CHARACTERIZATION AND IMPACT OF AMBIENT AIR POLLUTANT MEASUREMENT ERROR IN TIME-SERIES EPIDEMIOLOGIC STUDIES

A Dissertation
Presented to
The Academic Faculty

by

Gretchen Tanner Goldman

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy in the
School of Civil and Environmental Engineering

Georgia Institute of Technology
August 2011

CHARACTERIZATION AND IMPACT OF AMBIENT AIR POLLUTANT MEASUREMENT ERROR IN TIME-SERIES EPIDEMIOLOGIC STUDIES

Approved by:

Dr. James A. Mulholland, Advisor
School of Civil and Environmental Engineering
Georgia Institute of Technology

Dr. Armistead G. Russell
School of Civil and Environmental Engineering
Georgia Institute of Technology

Dr. Jian Luo
School of Civil and Environmental Engineering
Georgia Institute of Technology

Dr. Michael Chang
Brook Byers Institute for Sustainable Systems
Georgia Institute of Technology

Dr. Brian Stone, Jr.
School of City and Regional Planning
Georgia Institute of Technology

Dr. Matthew J. Strickland
Rollins School of Public Health
Emory University

Date Approved: June 7th, 2011

To my friends and family,
who have shown unwavering support for me throughout this effort

ACKNOWLEDGEMENTS

I would like to show gratitude for the many individuals who have helped make my Georgia Tech experience as great as it has been. These years down south have been a great adventure for me and I will forever look back on them with fondness. First and foremost, I would like to thank my primary advisor, Dr. James Mulholland. His unending guidance and open door policy has made for the best graduate experience I could have hoped for. His one-on-one academic support throughout my years at Georgia Tech, along with his willingness to let me explore peripheral opportunities at Tech (and openness to biking to Emory meetings with me in the Georgia summer heat!) have allowed me both personal and professional growth and I am truly grateful to have had the opportunity to work with him. I would also like to thank Dr. Ted Russell for his ongoing support of my work and willingness to devote his time to my projects whenever needed. His commitment to fostering a culture of mutual learning and shared resources among his students has greatly enhanced my academic experience. I would like to express my gratitude to Dr. Matthew Strickland for his work on the epidemiology side of my research. I am grateful for his open-minded and cooperative attitude in our ongoing discussion of new research ideas and this thesis would not be what it is, if not for his work on the project. Additionally, I would like to recognize the support of Drs. Paige Tolbert, Mitch Klein, Dana Flanders, Lance Waller and the rest of the SOPHIA research group at Emory University for their support and willingness to lend their expertise to my work. Their involvement in my research has proven invaluable and I appreciate the time and thought they have provided. I would like to acknowledge my other thesis committee

members, Dr. Jian Luo, Dr. Michael Chang and Dr. Brian Stone, Jr., for their valuable input on my research and continued support.

I want to thank my fellow students, whom have greatly enhanced my Georgia Tech and Atlanta experience. I am especially grateful to my peers, Jorge Pachon Quinche and Johana Husserl Orjuela, who have been right beside me through this process and have provided endless support in the form of homework help, qualifying exam study buddies, moral support in our mutual research trials and tribulations and celebration of milestones along the way. I am whole-heartedly appreciative of the friendship and support of Siv Balachandran, Farhan Akhtar, Burcak Kaynak, Jaemeen Baek, KJ Liao, Soonchul Kwon, Fernando Garcia Menendez, Aika Yano, Elaine Olivares and Radhika Dhingra and the rest of Dr. Russell and Dr. Mulholland's research group members whom I have had the privilege to work with. Further, I want to recognize my peers who have devoted considerable time and energies to teaching engineering concepts to a meteorologist during the coursework of my initial year at Georgia Tech. I am indebted to Abhishek Srivastava and Parker Johnson for their time and patience in explaining concepts and answering my questions. Additionally, I would like to thank Anita Gajjala, Anna-Sofia Kivelio, Lokesh Padhye, Samuel Huber, Jonathan Murphy, Liam Rattray and the countless other friends and colleagues who have helped shape my experience at Georgia Tech.

I cannot express enough thanks to my fiancé Tom for showing unyielding support of me following my dreams, even when they involved moving south of the Mason-Dixon line, years of Skype communication and listening to dozens of Powerpoint presentations on exposure measurement error. We have gone through this journey as a team and I know

I could not have done it without his support, both in Atlanta and from afar. Lastly, I would like to thank my family: Mom, Dad, Jared and Daniel, for their unconditional love and support of my academic and personal pursuits for as long as I can remember. From nursing me back to health after 4 sports-related surgeries to enduring 9-hour roundtrips to Ithaca, they have given me the ability to succeed and the confidence to fly. Throughout my life, they have allowed me the freedom to pursue my passions and I am certain I would not have made it this far without them.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Page

LIST OF SYMBOLS

h	distance between monitor observation locations
γ	raw semivariance
γ'	scaled semivariance
C	observed concentration in standard units
j, k	spatial location indices
Var	variance
i	day
R	Pearson correlation coefficient
γ'_o	nugget
γ'_e	partial sill
a_e	semivariogram parameter
r	distance from urban center
P	population density
A,B	population-weighting regression coefficients
R	radius of Atlanta study area
$\overline{\gamma}'$	population-weighted scaled semivariance
χ^*	normalized concentration
$\mu_{\ln C}$	time-series mean of log concentrations
$\sigma_{\ln C}$	time-series standard deviation of log concentrations
ε_χ	error in normalized concentrations
N	random number ~ N(0,1)

a,b	error model optimization constants
χ	simulated value with error added in normalized log concentration
Z	simulated value with error added in standard concentration units
σ_χ	standard deviation of χ
$\overline{R}$	Pearson correlation coefficient corresponding to $\overline{\gamma}'$
R_o	Pearson correlation coefficient corresponding to γ'_o
Y	count of emergency department visits
α	epidemiologic model y-intercept
β	log of the rate ratio for a change in pollutant concentration
$confounders$	vector of potential confounders
γ	vector of regression coefficients for suspected confounders
Z^*	true value (generic)
Z	measurement value (generic)
ε	error between Z and Z^* (generic)
β^*	true log of the rate ratio for a change in pollutant concentration
t	day
k,l	spatial locations
i,j	census tract indices
p	census tract population
p_{total}	total population of 20-county study area
m	first-order linear regression coefficient of Z-Z^* versus Z
IQR^*	true interquartile range
i	day
j	location

χ_{ij}^{*}	normalized "true" pollutant level on day *i* at location *j*
C_{ij}	concentration on day *i* at location *j*
$\mu_{\ln C_j}$	log concentration mean over all days at location *j*
$\sigma_{\ln C_j}$	log concentration standard deviation over all days at location *j*
Z_{ij}^{*}	"true" concentration on day *i* at location *j*
μ_r	log concentration mean modeled as a function of distance from urban center *r*
σ_r	log concentration standard deviation modeled as a function of *r*
α_{yr}	season factor modeled as a function of *r*
α_{wk}	day-of-week factor
Z	exposure metric
m	regressed slope of error versus measurement
b	regressed y-intercept of error versus measurement

LIST OF ABBREVIATIONS

ASACA	Assessment of Spatial Aerosol Composition in Atlanta
AQS	US EPA's Air Quality System
CI	confidence interval
CVD	cardiovascular disease
EC	elemental carbon
ED	emergency department visits
FRM	Federal Reference Method
IQR	interquartile range
km	kilometer
$\mu g/m^3$	micrograms per meter cubed
NH_4	ammonium
NO_3	nitrate
OC	organic carbon
PCM	particle composition monitor
PM	particulate matter
ppb	parts per billion
ppm	parts per million
R	Pearson correlation coefficient
RR	risk ratio
SEARCH	Southeastern Aerosol Research and Characterization Study
SO_4	sulfate
TEOM	tapered element oscillating microbalance
TOR	thermal optical reflectance

TOT	thermal optical transmittance

SUMMARY

Time-series studies of ambient air pollution and acute health outcomes utilize measurements from fixed outdoor monitoring sites to assess changes in pollution concentration relative to time-variable health outcome measures. These studies rely on measured concentrations as a surrogate for population exposure. The degree to which monitoring site measurements accurately represent true ambient concentrations is of interest from both an etiologic and regulatory perspective, since associations observed in time-series studies are used to inform health-based ambient air quality standards.

Air pollutant measurement errors associated with instrument precision and lack of spatial correlation between monitors vary widely across pollutants and these errors have been shown to attenuate associations observed in health studies. Further, the impact of measurement error varies depending on the type of error present, with classical error resulting in greater attenuation than Berkson error, which is expected to yield unbiased effect estimators. Characterization and adjustment for air pollution measurement error can improve effect estimates in time-series studies.

Measurement error due to instrument precision and spatial variability was characterized for ambient air pollutants in Atlanta. Error was modeled for daily measures of 12 air pollutants using measurements from collocated monitoring sites to characterize instrument precision and data from multiple study area monitors to estimate population-weighted spatial variance. Simulations of instrument and spatial error were generated for each pollutant, added to a reference pollutant time-series, and used in a Poisson generalized linear model of air pollution and cardiovascular emergency department visits. This method allows for pollutant-specific quantification of impacts of measurement error on health effect estimates, both the assessed strength of association and its significance.

The impact of these measurement error sources is affected by both the amount and the type of error. Regarding the latter, error simulations ranged in type from purely classical to purely Berkson, as defined on a log scale. To inform on the type and amount of error present in Atlanta measurements, air pollutant concentrations were simulated over the 20-county metropolitan area for a 6-year period, incorporating several distribution characteristics observed in measurement data. Spatial and temporal autocorrelation as well as trends for season, day-of-week and distance from downtown were modeled. The simulated concentration field was then used to characterize the amount and type of error due to spatial variability in ambient concentrations. The impact of use of different exposure metrics in a time-series epidemiologic study was assessed.

Finally, methodologies developed for the Atlanta area were applied to Dallas, Texas. Measurement error due to spatial variability was quantified for ambient monitoring site networks in Dallas with consideration for the impact of this error on a time-series study of Dallas that is currently underway. Differences in air pollution measurement error due to spatial variability between Atlanta and Dallas data were assessed and the impact of this measurement error on health associations in Dallas was discussed.

CHAPTER 1

INTRODUCTION

Air pollution has been linked to human health effects since the days of wood burning in ancient Rome (Hughes 1994). With the advent of the Industrial Revolution and the mass movement of people into cities, air pollution became a more widespread problem as a result of combustion emissions and exposure of larger populations (Jacobson 2002). While vast improvements have been made in our understanding and regulation of air pollution since the days of the London fog episode, air quality remains a valid public health concern as increases in urban populations and continued use of fossil fuels have allowed ambient air pollution to persist in many cities worldwide. In its assessment of the global burden of disease, the World Health Organization estimates ambient air pollution to cause approximately 2 million premature deaths globally on an annual basis (Cohen et al. 2005). Even in the United States where air pollution levels are comparatively low, ambient air pollutants have been linked to several adverse health outcomes including cardiorespiratory hospital admissions (Linn et al. 2000; Metzger et al. 2004; Moolgavkar et al. 1997), cardiorespiratory mortality (Hoek et al. 2001; Mar et al. 2000), lung cancer mortality (Abbey et al. 1999; Dockery et al. 1993) and adverse birth outcomes (Darrow et al. 2009; Hansen et al. 2003). In light of these linkages, much research continues to focus on understanding the association between air pollution and human health. While it is of value to understand the true causal relationship between exposure to air pollutants and health impact for etiologic or biological reasons, it is also of interest to understand the link between ambient pollutant levels and health impact from a regulatory standpoint as air pollution is regulated largely through use of fixed outdoor monitors.

Time-series studies utilize measurements from these outdoor monitoring stations to assess associations between ambient air pollution levels and time-variable health endpoints using epidemiologic regression models. In such studies, exposure measurement error is inherent as there is considerable uncertainty associated with air pollution measurement itself, as well as uncertainty in the degree to which measurements are representative of population exposure. Because evidence of associations between air pollution and health is drawn primarily from time-series studies and air quality regulation is largely health-based (Krewski and Rainham 2007), the direction and extent to which measurement error plays a role in these studies is of particular interest. Three components of exposure measurement error have been recognized: (1) the difference between individual exposures and average personal exposure, (2) the difference between average personal exposure and ambient levels and (3) the difference between measured ambient levels and true ambient concentrations (Zeger et al. 2000). This third component of measurement error can be further divided into two subcategories of (a) instrument error, which is inclusive of the accuracy and precision of monitoring instruments and (b) error resulting from the nonrepresentativeness of a monitoring site, which is reflected in the spatial variability of the pollutant measured (National Research Council 2001). The former, instrument error, can be attributed to instrument precision and drift effects, as well as any uncertainty resulting from sample preparation, collection, analysis, data acquisition, and data processing (International Organization for Standardization 1995). While instruments typically have reported uncertainty estimates, these estimates have been shown to underestimate uncertainties associated with actual use of the instrument when collocated instruments are used for uncertainty assessment (Hyslop and White 2008; White et al. 2005). Further, reported uncertainties may be inaccurate or otherwise unavailable (Dutton et al. 2009).

The latter source of error is any uncertainty resulting from the use of a single exposure metric to characterize ambient pollutant levels over a larger spatial area. Some

pollutant concentrations have been shown to vary widely within cities. For example, NO_2 concentrations have been shown to be spatially diverse even within 50 m distances (Hewitt 1991) and in an intra-city study of PM_{10} and $PM_{2.5}$, correlations between monitors were found to be as low as 0.4 (US Environmental Protection Agency 2009). In Atlanta, significant levels of spatial variability have been observed for several pollutants, particularly those emitted directly to the atmosphere from mobile or industrial sources such as NO_2, CO, SO_2, and the elemental carbon (EC) component of particulate matter less than 2.5 μm ($PM_{2.5}$) (Wade et al. 2006). Due to this observed spatial variability, prior studies have called for further investigations of spatial error and its correlation to measurements in cities with multiple monitors (Sheppard et al. 2005; Zeger et al. 2000).

Air pollution measurement error has been demonstrated to have significant impacts on time-series studies which often use measurements from a single central monitor to assess health endpoints over a broader spatial area. Sheppard et al. (2005) found that measurement error resulting from spatial variability led to an attenuation of the health effect parameter β of 7.7% for $PM_{2.5}$ mass when exposure was estimated from a single monitor in an ecological time-series study. In a daily time-series study of ambient $PM_{2.5}$ mass and cardiovascular mortality, observed associations were found to be lower in geographic subpopulations residing farther from a central monitoring site compared to populations in closer proximity to the central site (Wilson et al. 2007). In an assessment of spatial error in Atlanta, significant health associations for more spatially homogeneous $PM_{2.5}$ and O_3 and emergency department visits were observed regardless of the spatial location of the monitor used for assessment, while for more spatially heterogeneous pollutants, NO_2 and CO, significant associations were found when measurements from urban monitoring sites were used for assessment; however, no association was observed when a rural monitoring site was used (Sarnat et al. 2010). These results suggest that significance of association may be influenced by the spatial variability of the metric measured and that air pollution measurement error resulting from the spatial

heterogeneity of ambient concentrations can significantly bias effect estimates in time-series studies.

This observed bias in health associations due to measurement error has been shown to vary based on the error type present (Armstrong 1998; Zeger et al. 2000). While error can fall anywhere along a continuum in the conceptual framework of error type, two distinctly different types, classical error and Berkson error, have been recognized. Under a classical error framework, the measurement, Z, is an imperfect measure of the true exposure, Z^*, and as a result, measurements vary randomly about the truth, that is $E[Z|Z^*] = Z^*$. Because this type of error adds noise to the independent variable in the health outcome regression, an attenuation of the effect estimator, β, occurs. A distinctly different error type is Berkson error. Under a Berkson error framework, the measurement, Z, is the average value of true exposures, Z^*, which vary about Z, that is, $E[Z^*|Z] = Z$. Individual true exposures vary randomly about the measurement; thus, the average measured exposure is an unbiased estimate of total exposure and although a purely Berkson type error will add uncertainty to the health association by way of decreased significance of association, no attenuation of the effect point estimate is expected.

Measurement error due to spatial variability has been estimated to be largely Berkson in nature, with the true exposures of individuals varying randomly about some measured population average exposure (Zeger et al. 2000). Instrument error might be thought of as classical error with individual measurements of an instrument equally likely to measure high or low relative to the true concentration which is measured by the instrument on average (Goldman et al. 2011). Although error type has been estimated in these and other studies, because knowledge of the type of error present in a given set of measurements is reliant upon knowledge of true ambient concentrations which cannot be known, assessment of error type and adjustment for it in health studies remains a prominent challenge in time-series studies of ambient air pollution.

Several studies have used geostatistical simulation of air pollution to characterize ambient variability in concentrations and assess its impact in time-series studies (Lee and Shaddick 2010; Peng and Bell 2010). Such simulation studies can provide insight into the amount and type of error attributable to spatial variability in ambient levels.

It was hypothesized that both measurement error amount and type impact the bias observed in health associations in time-series epidemiologic studies, with error amount due to spatial variability resulting in greater attenuation of risk estimates than error amount due to instrument precision and classical error resulting in greater attenuations than Berkson error. The approach of this research was (i) to characterize ambient air pollution measurement error due to instrument precision and spatial variability and assess their impact on a time-series epidemiologic study, (ii) to examine how the impact of measurement error is modified by error type, (iii) to determine the amount and type of error present in an ongoing epidemiologic study through geostatistical simulation and (iv) to utilize methods developed for measurement error characterization and impact assessment in an Atlanta study to estimate measurement error impacts for an epidemiologic study in Dallas, Texas.

CHAPTER 2

OBJECTIVES

The objective of this dissertation was to characterize air pollution measurement error in large scale time-series epidemiologic studies due to instrument error and spatial variability and heterogeneity of ambient concentrations, through analysis of measurements and simulation of air pollution concentrations. Given below are objectives for individual chapters.

Chapter 3: Ambient Air Pollutant Measurement Error: Characterization and Impacts in a Time-Series Epidemiologic Study in Atlanta

Chapter 3 describes a study in which air pollution measurement error due to instrument precision and spatial variability of ambient pollutants is characterized using measurements from collocated instruments and geostatistical semivariogram analysis, respectively, for 12 air pollutants. The impact of these error sources on a time-series epidemiologic study in Atlanta is quantified. This work has been published (Goldman et al. 2010).

Chapter 4: Impact of Exposure Measurement Error in Air Pollution Epidemiology: Effect of Error Type in Time-Series Studies

Chapter 4 describes a study of the impact of error type on the assessment of measurement error in a time-series epidemiologic study in Atlanta. Simulated air pollution time-series are generated over a range of error types between classical and Berkson as defined on a log scale and the corresponding range of impacts on a health association is estimated. This study takes a comprehensive look at the impact of error type on time-series studies of air pollution. This work has been published (Goldman et al. 2011).

Chapter 5: Measurement Error in Time-series Studies of Ambient Air Pollution: A Simulation Study

Chapter 5 describes a simulation study in which a spatio-temporal model is developed to generate ambient air pollutant concentrations over a 20-county area and 6-year time period. Several distribution characteristics of measurement data are modeled, including spatial and temporal autocorrelation of values. The concentration field is utilized to assess the amount and type of error associated with spatial availability of ambient concentration and use of monitor-based exposure metrics in time-series studies. A manuscript is in preparation.

Chapter 6: Spatio-temporal Variability of Ambient Air Pollutants in Dallas, Texas: Implications for Time-Series Epidemiologic Studies

Chapter 6 describes a study exploring the spatio-temporal distribution of ambient air pollutants in Dallas, Texas. The methodology for characterization of measurement error due to spatial variability developed for Atlanta is applied to monitoring site data for Dallas. Measurement error due to spatial variability is assessed and the potential impact of this error on a planned time-series health study is discussed.

Chapter 7: Conclusions and Future Work

Chapter 7 provides conclusions of the dissertation and offers recommendations on future directions for this work.

CHAPTER 3

AMBIENT AIR POLLUTANT MEASUREMENT ERROR: CHARACTERIZATION AND IMPACTS IN A TIME-SERIES EPIDEMIOLOGIC STUDY IN ATLANTA

(Goldman GT, Mulholland JA, Russell AG, Srivastava A, Strickland MJ, Klein M, Waller LA, Tolbert PE, and Edgerton ES. *Environmental Science & Technology*, 44: 7692-7698, 2010)

Abstract

In time-series studies of ambient air pollution and health in large urban areas, measurement errors associated with instrument precision and spatial variability vary widely across pollutants. In this paper, we characterize these errors for selected air pollutants and estimate their impacts on epidemiologic results from an ongoing study of air pollution and emergency department visits in Atlanta. Error was modeled for daily measures of 12 air pollutants using collocated monitor data to characterize instrument precision and data from multiple study area monitors to estimate population-weighted spatial variance. Time-series simulations of instrument and spatial error were generated for each pollutant, added to a reference pollutant time-series, and used in a Poisson generalized linear model of air pollution and cardiovascular emergency department visits. Reductions in risk ratio due to instrument precision error were less than 6%. Error due to spatial variability resulted in average risk ratio reductions of less than 16% for secondary pollutants (O_3, $PM_{2.5}$ sulfate, nitrate and ammonium) and between 43% and 68% for primary pollutants (NO_x, NO_2, SO_2, CO, $PM_{2.5}$ elemental carbon); pollutants of mixed

origin (PM_{10}, $PM_{2.5}$, $PM_{2.5}$ organic carbon) had intermediate impacts. Quantifying impacts of measurement error on health effect estimates improves interpretation across ambient pollutants.

3.1 Introduction

Exposure measurement error is inherent to time-series studies of ambient air pollution and health in large populations and its assessment has been cited as a research priority(Brauer et al. 2003; Brunekreef and Holgate 2002; Jerrett et al. 2005; National Research Council 2001; US Environmental Protection Agency 2009). Zeger et al. (2000) describe three components of measurement error: (1) differences between individual exposures and average personal exposure, (2) differences between average personal exposure and ambient levels, and (3) differences between measured and true ambient concentrations. In this paper, we investigate this third component of measurement error, which can be divided into two subcategories (National Research Council 1998): (a) instrument error, which results from inaccuracy and imprecision in the ambient monitor observations, and (b) spatial error, which results from the inability of a single time-series to accurately represent the ambient levels throughout the study area.

Instrument error can be examined through analysis of measurements from collocated monitors. Discrepancy between independent measurements at the same location may be due to instrument calibration and drift, as well as errors in sample preparation, collection, analysis, data acquisition, and data processing. Uncertainty estimates based on laboratory tests often underestimate error associated with actual use of the instrument when assessed from collocated instrument data (Hyslop and White 2008; White et al. 2005).

A second source of error addressed here is that resulting from the use of a single measure (e.g., central monitor data or spatial average estimates using data from multiple monitors) to characterize ambient pollutant levels over a study area. Both micro- and

macro-scale variations in ambient air pollution have been observed in metropolitan areas (Hewitt 1991). Spatial variation, if not independent of ambient concentration, can bias effect estimates in time-series health studies. Sheppard et al. (2005) found that measurement error resulting from spatial variability led to an attenuation of acute health effect estimates of 7.7% for $PM_{2.5}$ mass when exposure was estimated from a single monitor. Wilson et al. (2007) observed associations between ambient $PM_{2.5}$ and cardiovascular mortality to be lower in geographical subpopulations farther from a central monitoring site compared with populations close to the central site in a daily time-series study. While investigating three spatial models in a study of airborne particles and respiratory emergency hospital admissions, Chen et al. (2007) found that effect estimates of PM_{10} were underestimated when a non-spatial approach was used. In a recent study of ambient air pollution and emergency department visits, effect estimates for spatially heterogeneous pollutants (CO and NO_2) were dependent on which monitor was used, whereas observed associations for spatially homogeneous pollutants (O_3 and $PM_{2.5}$) were similar regardless of the monitoring site (Sarnat et al. 2010). These results suggest that there may be greater attenuation of health risk estimates for spatially heterogeneous primary pollutants than for spatially homogeneous secondary pollutants, as one might expect.

In this paper, we address the need for a comprehensive examination of measurement error impacts on health risk estimates in a study of air pollution and emergency department (ED) visits for cardiovascular disease (CVD) in Atlanta. Our objectives were, first, to characterize and model measurement error using ambient air monitoring data, and, second, to assess the impact of measurement error by rerunning epidemiologic models using reference pollutant data with modeled error added. For selected ambient air pollutants in the metropolitan Atlanta region, we provide a detailed characterization of both instrument precision error and error due to spatial variability as these errors have been previously demonstrated to be the relevant sources of potential

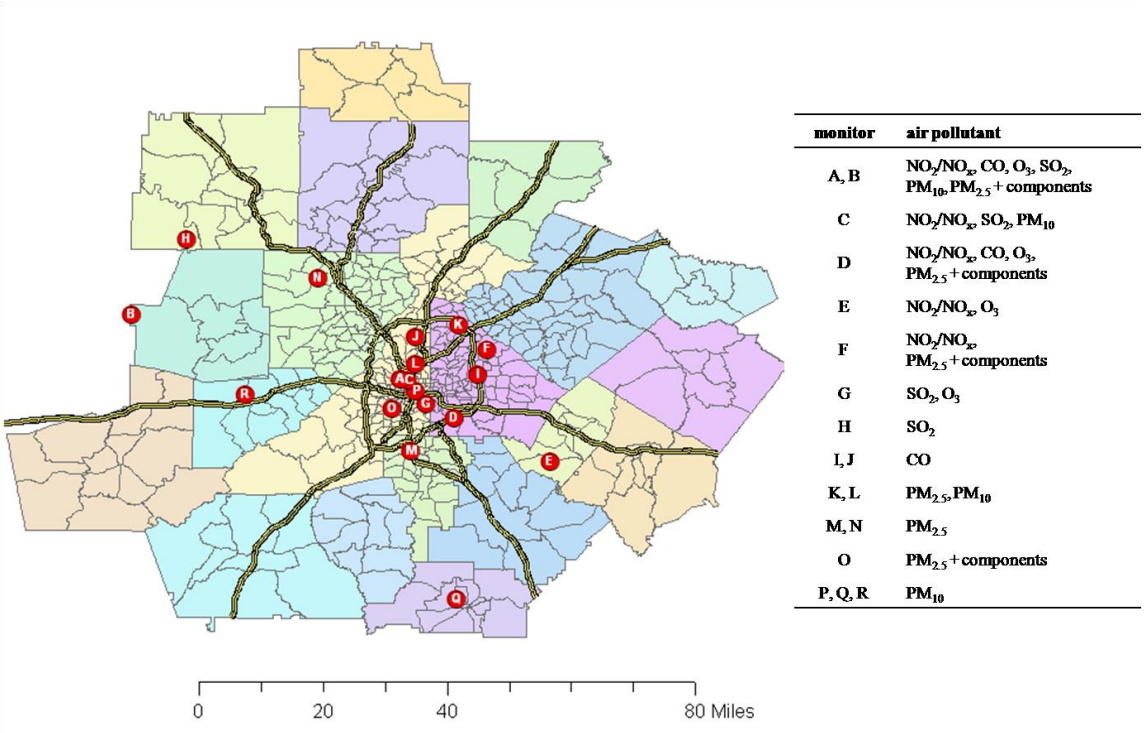
bias in regression coefficients in time-series studies (Carrothers and Evans 2000). This study is limited to variability in ambient air pollution as assessed from regulatory and regional study monitors at fixed locations throughout the area of interest; micro-scale variability, such as that associated with near-roadway concentrations, is not addressed; nor is temporal micro-scale variability addressed, such as that associated with meteorological events on sub-hour time scales. Micro-scale spatial and temporal variability in pollutant levels may be sources of error in etiologic investigations of acute health effects from air pollution, but this study examines the impact of error on health risk estimates for ambient levels of pollutants measured in accordance with regulatory specifications. Finally, this study addresses the effects of instrument imprecision and spatial variability, not the effects of instrument inaccuracy and spatial heterogeneity (e.g. urban-rural differences in primary air pollutant concentrations). The former result in day-to-day lack of correlation between air pollutant measurements at either the same location with different instruments or at different locations; the latter reflect systematic differences in pollutant concentration measurements between instruments or over space.

3.2 Methods

Our approach for characterizing measurement error and assessing its impact in a time-series health study involves five steps: (1) collection of ambient air pollutant data; (2) characterization of instrument precision by analyzing collocated instrument data; (3) characterization of spatial variability using geostatistical methods; (4) generation of simulations by adding modeled error to reference pollutant data; (5) use of the simulations in epidemiologic models and comparison of results with those obtained using the reference time-series data. Methods used in each of these steps follow.

3.2.1 Ambient Air Pollutant Data

Twelve pollutants were assessed: NO_2, NO_x, O_3, SO_2, CO, PM_{10} mass, $PM_{2.5}$ mass, and $PM_{2.5}$ components sulfate (SO_4), nitrate (NO_3), ammonium (NH_4), elemental carbon (EC) and organic carbon (OC). Metropolitan Atlanta area monitors are shown in Figure 3.1. Air pollution measurements from three sources were utilized: the US EPA's Air Quality System (AQS), including State and Local Air Monitoring System and Speciation Trends Network for $PM_{2.5}$ component measurements; the Southeastern Aerosol Research and Characterization Study (SEARCH) network (Hansen et al. 2003), including the Atlanta EPA supersite at Jefferson Street (Solomon et al. 2003b)**;** and the Assessment of Spatial Aerosol Composition in Atlanta (ASACA) network (Butler et al. 2003). To assess measurement error due to instrument imprecision and spatial variability, 1999-2004 datasets were used, except where noted.



monitor	air pollutant
A, B	NO_2/NO_x, CO, O_3, SO_2, PM_{10}, $PM_{2.5}$ + components
C	NO_2/NO_x, SO_2, PM_{10}
D	NO_2/NO_x, CO, O_3, $PM_{2.5}$ + components
E	NO_2/NO_x, O_3
F	NO_2/NO_x, $PM_{2.5}$ + components
G	SO_2, O_3
H	SO_2
I, J	CO
K, L	$PM_{2.5}$, PM_{10}
M, N	$PM_{2.5}$
O	$PM_{2.5}$ + components
P, Q, R	PM_{10}

Figure 3.1. Map of 20-county metropolitan Atlanta study area. Census tracts, interstate highways, and ambient air pollutant monitoring sites are shown.

For several pollutants, different measurement methods were used. SEARCH monitors (sites A and B) independently measure NO and NO_2, whereas the AQS NO_x measurement may include additional oxides of nitrogen. For SO_2 measurement, less SO_2 is lost by water condensation in the SEARCH sampling system than in the AQS measurement. For $PM_{2.5}$ mass measurements, both Federal Reference Method (FRM) instruments, which provide 24-hr average filter-based measurements, and Tapered Element Oscillating Microbalance (TEOM) instruments, which provide 1-hr semi-continuous data, were used. The TEOM analyzers were operated to minimize loss of semi-volatiles (Edgerton et al. 2006). For $PM_{2.5}$ components, data were obtained from particle composition monitors (PCM) providing 24-hr filter samples, except for collocated data which were obtained using continuous methods. Ions (SO_4, NO_3, NH_4) were detected by chromatographic analysis; for carbon data, SEARCH uses the thermal optical reflectance method for differentiating EC and OC, whereas AQS uses the thermal optical transmittance method. The EC-OC split differs between these methods (Chow et al. 2001); however, the measures are highly correlated (Solomon et al. 2003a). The differences between measurement methods are expected to have little impact on the assessment of instrument error in this study because error here is based on the correlation, not the bias, between measurements.

3.2.2 Instrument Precision

Collocated instrument data during 1999-2004 were available for NO_2, NO_x and O_3 at Yorkville (Figure 3.1, site B) from AQS and SEARCH monitors. For $PM_{2.5}$ total mass and major components, collocated filter-based measurements and unadjusted continuous data (Edgerton et al. 2006) at the Jefferson Street SEARCH site (Figure 3.1, site A) were used over the six-year period 2002-2007. For SO_2, CO and PM_{10}, collocated instrument data were not available.

Continuous $PM_{2.5}$ ion measurements tend to underestimate concentrations (Chow et al. 2008; Hogrefe et al. 2004), which has been attributed to excess water in the instrument and reductions in flash volatilization efficiency at high concentrations (Chow et al. 2008; Long and McClenny 2006). For carbon fractions, the filter-based EC and OC measurements by TOR analysis were compared with aethalometer measurement of black carbon and a semi-continuous total carbon measurement with the aethalometer black carbon subtracted, respectively. Compounded error in the continuous OC estimates is expected to result in conservatively high estimates of instrument error for this pollutant.

3.2.3 Spatial Variability

The semivariogram provides information on the spatial correlation of data and has been used previously for air pollutant data (Casado et al. 1994; Wade et al. 2006). In this study, modified semivariograms were used to assess spatial variability error over the entire six-year study period. In the raw semivariogram, $\gamma(h)$, defined here as one-half of the variance of the differences between observations (C_j and C_k) at two locations (j and k) located a distance h apart over all days, is plotted versus h. Here, the semivariance is normalized by the temporal variance of the average of two observations to yield a scaled semivariance, γ'.

$$\gamma'(h) \equiv \frac{Var\left(\frac{C_j - C_k}{2}\right)}{Var\left(\frac{C_j + C_k}{2}\right)} = \frac{\text{spatial semivariance}}{\text{temporal variance}} \tag{3.1}$$

Thus, γ' represents the spatial semivariance scaled to a quantity indicative of the range of exposures over which health risk is being assessed; it is unitless and allows for comparison across pollutants. It can be shown that this scaled semivariance is related to the Pearson correlation coefficient (R) between observations at two sites as follows (Wade et al. 2006).

$$\gamma'(h) = \frac{1 - R(h)}{1 + R(h)} \tag{3.2}$$

Thus, a scaled semivariance value of 0 corresponds to perfectly correlated observations (R = 1) and a value of 1 corresponds to perfectly uncorrelated observations (R = 0). The scaled semivariance as defined by eqs. 3.1 and 3.2 is, therefore, a measure of spatial variability for the entire study period that can be compared across pollutants. Since the power to observe an association in a time-series health study depends on the temporal variability of pollutant concentrations, the impact of error due to spatial variability will depend on how large the spatial variance is relative to the temporal variance, evaluated for the entire study population over the entire study period.

Assuming the spatial variation of air pollutants to be isotropic, the scaled semivariance was modeled as a function of the distance between observations, *h*, as follows.

$$\gamma'(h) = \gamma'_o + \gamma'_e \left\{1 - \exp\left(-\frac{h}{a_e}\right)\right\} \tag{3.3}$$

Here, γ'_o, called the nugget, is the semivariance when monitors are collocated and, therefore, represents instrument precision error scaled by the temporal variance. Collocated monitor data were analyzed to determine the nugget semivariance and, thus, characterize instrument precision error. The sill, $\gamma'_o + \gamma'_e$, is the asymptote of the semivariance, and the range, $3a_e$, is the distance at which 95% of the sill is reached. The exponential model above provided a reasonable fit to the data, as will be shown.

We used the theoretical semivariogram to compute an average scaled semivariance relative to the urban center for the entire study population for each pollutant. Data from the 2000 census for 660 tracts in the metropolitan Atlanta study area were used to develop the following relationship between population density, *P(r)*, and distance from the urban center, *r* (Figure A.1).

$$P(r) = Ae^{-Br} \tag{3.4}$$

Parameters A and B are regression coefficients. This relationship (eq. 3.4) was used with the modeled semivariogram (eq. 3.3) to calculate an integrated population-weighted semivariance, $\overline{\gamma'}$, for each pollutant.

$$\overline{\gamma'} = \frac{\int_0^R \gamma'(r)P(r)rdr}{\int_0^R P(r)rdr} = \gamma'_o + \gamma'_e \left\{1 - \left(\frac{B}{B + a_e^{-1}}\right)^2 \left(\frac{\left(BR + a_e^{-1}R + 1\right)e^{-(B+a_e^{-1})R} - 1}{(BR+1)e^{-BR} - 1}\right)\right\} \tag{3.5}$$

Here, the radius of the Atlanta study area, R, is 80 km, B is 0.052 km^{-1}; the nugget (γ'_o), partial sill (γ'_e) and range ($3a_e$) are determined for each pollutant (Table A.1). Since a non-zero nugget is included in the integration, the population-weighted semivariance includes instrument imprecision effects. We hypothesize that this integrated scaled semivariance can be used to predict the impact of error due to spatial variability on health risk estimation.

3.2.4 Simulations

To develop simulated datasets with modeled instrument and spatial error added to measurements, the steps shown in Figure 3.2 were taken. First, a reference time-series was chosen so that impacts on health effect estimates of adding error for each pollutant could be compared to the same base case. Based on preliminary epidemiologic model results, we chose the 1-hr maximum CO time-series at a central monitor (site A) as our base case because of its significant positive association with ED visits for CVD. Site A was chosen as the base case because it is a former EPA supersite and is heavily used in epidemiologic studies in Atlanta. Second, the log-normally distributed CO concentrations were normalized to avoid negative concentrations and preserve log-normality after error is added.

$$\chi_i^* = \frac{\ln C_i - \mu_{\ln C}}{\sigma_{\ln C}} \tag{3.6}$$

Here, χ_i^* is the normalized concentration on day *i* and $\mu_{\ln C}$ and $\sigma_{\ln C}$ are the mean and standard deviation, respectively, of the log concentrations over all days; thus, the mean and standard deviation of χ_i^* are 0 and 1, respectively. Third, error in the normalized concentration was modeled as a function of concentration based on the observed dependence of error on pollutant concentration.

$$\varepsilon_{\chi i} = N_i a C_i^b \tag{3.7}$$

Here, $\varepsilon_{\chi i}$ is the modeled error in χ_i^* for day *i*, N_i is a random number with distribution $\sim N(0,1)$ and inclusive of short-term temporal autocorrelation, and *a* and *b* are constants obtained by optimization under the constraints described below. Fourth, the simulated error was introduced to normalized data (eq. 3.8), and, fifth, to provide simulations of CO data with error added that have similar distributions to the original CO data, these were denormalized (eq. 3.9).

$$\chi_i = \chi_i^* + \varepsilon_{\chi i} \tag{3.8}$$

$$Z_i = \exp\left(\chi_i \frac{\sigma_{\ln C}}{\sigma_\chi} + \mu_{\ln C} \right) \tag{3.9}$$

Here, σ_χ is the standard deviation of χ_i, which is slightly larger than 1 due to the error added to χ_i^*. Thus, the simulations with error added (Z_i) have the same log mean and standard deviation as the central monitor CO data. These simulations represent alternative time-series derived from the same base case (i.e. the "true" time-series, Z^*) with error added and with similar distributions as the base case.

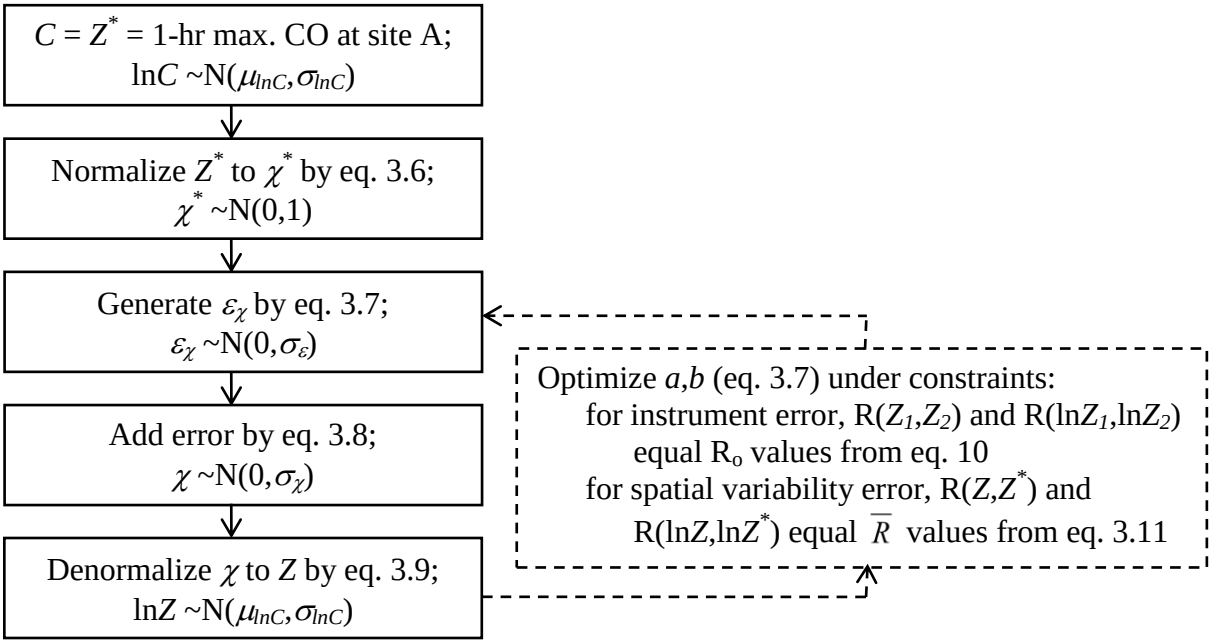


Figure 3.2. Time-series simulation flow chart. For each error type and each pollutant, the procedure was repeated 20 times to obtain average of *a* and *b*, then repeated without optimization (dashed lines) to generate 1000 simulated time-series.

For each pollutant, two semivariograms were constructed, one using log-concentration data and the other using concentration data directly. The corresponding two semivariogram nuggets and two population-weighted semivariances provide constraints for optimizing *a* and *b* for instrument imprecision and spatial variability, respectively. These constraints are calculated from the semivariogram nuggets and population-weighted semivariances by inverting eq. 3.2, as shown in eq. 3.10 and eq. 3.11, respectively.

$$R_o = \frac{1-\gamma'_o}{1+\gamma'_o} \tag{3.10}$$

$$\overline{R} = \frac{1-\overline{\gamma'}}{1+\overline{\gamma'}} \tag{3.11}$$

For instrument error, since the relationship between measurements is known but the relationship between measurement and truth is not, simulated time-series datasets are constrained to have inter-correlations, R(lnZ_1,lnZ_2) and R(Z_1,Z_2), that equal those from

collocated instruments (R_o). It was found that these constraints yielded error of the same magnitude and concentration dependence as observed in the data (Figure A.2). For spatial error, defined here as the error that results from extrapolating central site exposures to locations removed from the central site, simulated time-series datasets are constrained to have correlations with the base case, $R(\ln Z, \ln Z^*)$ and $R(Z, Z^*)$, that equal the average of correlations of ambient pollution time-series datasets at residences of all people in the study area with a central ambient pollutant time-series ($\overline{R}$). This optimization procedure was repeated 20 times to obtain average *a* and *b* values for use in 1000 Monte Carlo simulations.

Three measurement error models were developed for each of 12 pollutants in order to simulate three scenarios: (1) instrument error with semivariogram nuggets constrained and temporal autocorrelation of error included; (2) spatial error with population-weighted semivariances constrained and temporal autocorrelation of error included; (3) spatial error simulations with population-weighted semivariances constrained but without temporal autocorrelation of error included. Thus, 36 sets of 1000 simulated time-series for the six-year period 1999-2004 of central monitor (site A) CO data with error added were produced.

3.2.5 Epidemiologic Model

Relationships between daily measures of ambient air pollution and daily counts of ED visits for CVD (including ischemic heart disease, dysrhythmia, congestive heart failure, and peripheral/cerebrovascular disease) were assessed using methods described elsewhere (Metzger et al. 2004) and briefly summarized here. There were 166,950 ED visits for CVD in the 20-county metropolitan Atlanta area during 1999-2004. Lag 0 associations between daily pollutant concentration and the daily count of ED visits were assessed using Poisson generalized linear models that accounted for overdispersion. The general form of the epidemiologic model is

$$\log[E(Y_i)] = \alpha + \beta\ Z_i + \gamma\ confounders_i \quad (3.12)$$

where Y_i is the count of emergency department visits on day i, Z_i is the pollutant concentration on day i at a central monitor, and $confounders_i$ is the vector of potential confounders on day i. The specific potential confounders included in the model were indicator variables for day-of-week, season, and when a hospital entered or left the study; cubic terms for maximum temperature and dew point; and a cubic spline with monthly knots for day of follow-up. The parameter α is the intercept, β is the log of the rate ratio for a change in pollutant concentration, and γ is the vector of regression coefficients for the suspected confounders included in the model. The risk ratio (RR) is given by eq. 3.13.

$$RR = e^{\beta} \quad (3.13)$$

Using data from the central monitor (site A), preliminary epidemiologic assessments were performed for all air pollutants and ED visits for CVD. Consistent with previous findings (Metzger et al. 2004), significant positive associations were found for several traffic-related pollutants, including 1-hr maximum NO_x, 1-hr maximum CO and 24-hr average EC. To compare the impacts of adding error across all pollutants, we selected 1-hr maximum CO for our base case time-series. Monte Carlo simulations of instrument and spatial error added for each pollutant to the base case were used to test the impact of measurement error on epidemiologic analyses. With this approach, the actual central monitor CO observations are considered the "true" values (Z^*), and the simulated datasets with error added are the "measured" values (Z). A total of 36,000 epidemiologic analyses were performed using simulated time-series.

3.3 Results and Discussion

3.3.1 Ambient Air Quality Data

Daily metrics for 12 ambient air pollutants were calculated: 1-hr maximum NO_2, NO_x, SO_2 and CO, 8-hr maximum O_3, and 24-hr average PM_{10}, $PM_{2.5}$ and $PM_{2.5}$

components NO_3, SO_4, NH_4, EC and OC. These measures were found to be best described by lognormal distributions. Lognormal distribution properties (geometric mean, μ_g, and standard deviation, σ_g) for all 12 pollutants measured at the central SEARCH monitor (site A) and at the rural SEARCH monitor (site B) are listed in Table 3.1. Data completeness for 1999-2004 (2,192 days) ranged from 82% to 97% across the pollutants. Levels of primary pollutants (NO_2, NO_x, SO_2, CO and EC) are much higher at the urban monitor and much less spatially correlated (urban-rural R ranging from 0.10 to 0.43). For pollutants of largely secondary origin (O_3, NO_3, SO_4 and NH_4), the average levels are similar at urban and rural sites and the spatial correlation is much higher (urban-rural R ranging from 0.72 to 0.92). Pollutants of mixed origin ($PM_{2.5}$, PM_{10} and OC) have an intermediate level of spatial heterogeneity and spatial correlation.

Table 3.1. Lognormal distribution properties of urban (site A) and rural (site B) SEARCH monitors, 1999-2004; collocated instrument and urban-rural spatial Pearson correlation coefficients are also shown.

	urban (site A)		rural (site B)		collocated	urban-rural
	μ_g	σ_g	μ_g	σ_g	R	R
1-hr max NO_2	38.1 ppb	1.51	7.74 ppb	2.47	0.92 [a]	0.10
1-hr max NO_x	86.8 ppb	2.23	8.39 ppb	2.56	0.94 [a]	0.21
8-hr max O_3	35.8 ppb	1.92	45.2 ppb	1.54	0.99 [a]	0.89
1-hr max SO_2	11.4 ppb	2.58	6.32 ppb	2.40	N/A	0.10
1-hr max CO	0.89 ppm	2.11	0.24 ppm	1.39	N/A	0.18
24-hr PM_{10}	23.8 $\mu g/m^3$	1.55	17.2 $\mu g/m^3$	1.68	N/A	0.78
24-hr $PM_{2.5}$	15.2 $\mu g/m^3$	1.62	11.5 $\mu g/m^3$	1.73	0. 97 [b] (0.991) [c]	0.85
24-hr $PM_{2.5}$-SO_4	3.85 $\mu g/m^3$	1.95	3.71 $\mu g/m^3$	2.07	0. 96 [b] (0.998) [d]	0.92
24-hr $PM_{2.5}$-NO_3	0.74 $\mu g/m^3$	2.13	0.63 $\mu g/m^3$	2.20	0. 94 [b] (0.993) [d]	0.80
24-hr $PM_{2.5}$-NH_4	2.02 $\mu g/m^3$	1.79	1.97 $\mu g/m^3$	1.95	0. 91 [b] (0.994) [d]	0.72
24-hr $PM_{2.5}$-EC	1.28 $\mu g/m^3$	1.87	0.57 $\mu g/m^3$	1.81	0. 92 [b] (0.903) [d]	0.43
24-hr $PM_{2.5}$-OC	3.79 $\mu g/m^3$	1.65	2.70 $\mu g/m^3$	1.71	0.85 [b] (0.979) [d]	0.59

[a] Yorkville (site B), SEARCH and AQS monitors, 1999-2004
[b] Jefferson St (site A), filter-based and continuous monitors, 2002-2007
[c] Yorkville (site B), two filter-based FRM instruments, 2002-2008
[d] Centreville, Alabama, two filter-based instruments, 2001-2008

3.3.2 Instrument Precision Error

The collocated instrument correlations in Table 3.1 are indicative of the amount of instrument error for each pollutant. For $PM_{2.5}$ total mass and major constituents, the correlations between filter-based measurements and measurements using continuous methods are lower than correlations between collocated data from a Centreville, Alabama site using identical filter-based instruments (shown in parentheses). Thus, our instrument error estimations are conservative in that they include some variability due to different instruments or analytical methods. The collocated OC observations were least correlated, likely due to compounded error in the continuous OC measurement involving taking the difference between total carbon and black carbon observations.

Collocated instrument time-series were used to calculate nugget semivariance values on a concentration basis and a log-concentration basis, which characterize instrument precision in the error simulation model. In the case of CO, SO_2 and PM_{10} for which we did not have collocated instrument data, a value fitted by the semivariogram was used. All nugget semivariance values are listed in Table A.1.

In general, instrument error was observed to increase with increasing concentration; this concentration dependence was modeled via eq. 3.7. Short-term temporal autocorrelation of error was also observed, with correlation coefficients averaged across all pollutants of 0.59 ± 0.13 (standard deviation) and 0.33 ± 0.17 for one-day and two-day lags, respectively; this was modeled by using a three-day running average of random numbers for N_i (eq. 3.7). The desired level of error autocorrelation was verified in all models, as was the concentration dependence of the magnitude of error. The most extreme outliers in the observed error were not captured, however (Figure A.2).

3.3.3 Spatial Variability Error

Correlations between observations from all pairs of monitors measuring the same pollutant during 1999-2004 were calculated on a concentration basis and log concentration basis. As shown in Figure 3.1, there were six NO_2 and NO_x monitoring sites, five sites each for SO_2, CO and O_3 (with three O_3 sites providing only March-November data), eight PM_{10} sites (with five providing data only every 6 days), nine $PM_{2.5}$ mass sites (with four providing data only every 3 days), and five sites with $PM_{2.5}$ composition monitors. Scaled semivariograms were constructed for both concentrations and log concentrations, using eq. 3.2, and plotted as a function of distance between monitors. In Figure 3.3, log concentration semivariograms are shown for all pollutants. Exponential theoretical semivariograms were fit to the data (eq. 3.3) by using least squares regression to determine the range, assuming a sill of 1 and using a nugget derived from the instrument precision analysis.

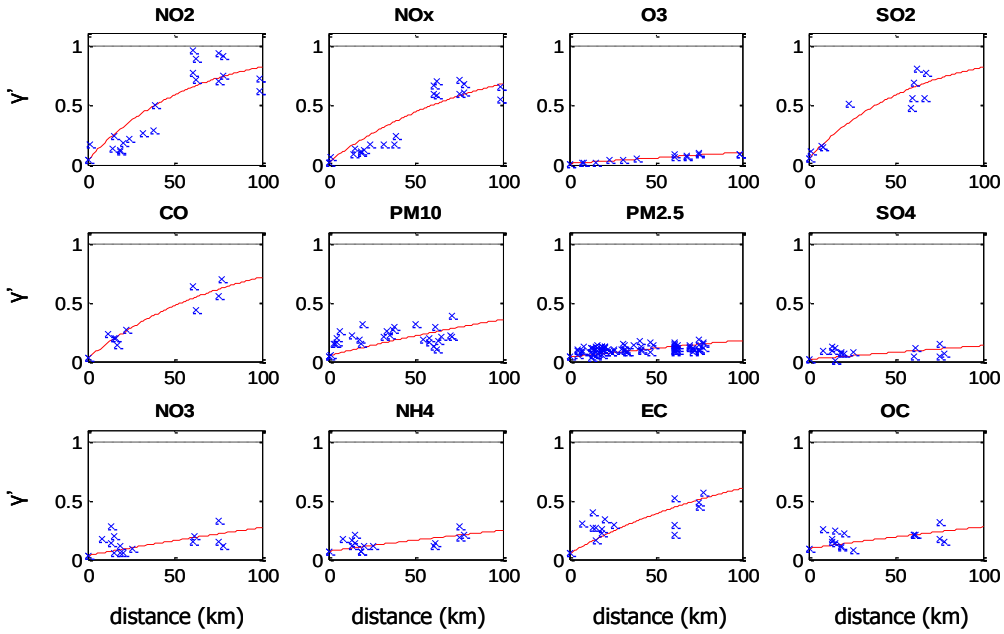


Figure 3.3. Semivariograms for the log normalized pollutant concentrations. Exponential curves are fitted to the data. Similar semivariograms were constructed for normalized pollutant concentrations.

Secondary pollutants (O_3, NO_3, SO_4 and NH_4) have large range values (from 1100 km to 3000 km), indicating their greater spatial autocorrelation. Primary pollutants (NO_2, NO_x, SO_2, CO and EC), conversely, have lower range values (from 100 to 330 km), demonstrating greater spatial variability of these pollutants. Scatter in the raw semivariogram suggests a degree of anisotropic behavior (i.e. directional dependence) in the data in addition to observational error.

Population-weighted semivariances were derived from both the concentration and log concentration semivariograms (Table A.1). While nuggets were all small (<0.1), indicative of low levels of instrument error, population-weighted semivariograms suggest much larger error due to spatial variability. As mentioned previously, secondary pollutants exhibit much less spatial variability than primary pollutants. Ozone exhibits the least spatial variability, and SO_2 the most. Moreover, SO_2 spatial variability is likely underestimated due to an insufficient number of monitors for this pollutant whose emissions are dominated by a few point sources.

3.3.4 Simulations with Modeled Error Added

Optimization yielded values of parameters *a* and *b* that were very similar across the 10 trials (Table A.2) for each error type and each pollutant, indicative of a relatively stable solution to the error model optimization routine. After optimization of model parameters *a* and *b*, 1000 simulations were generated for each pollutant by adding instrument precision error and error due to spatial variability, both with and without temporal autocorrelation, to the reference time-series via the procedure outlined in Figure 3.2. Average correlation coefficients between simulations agreed with expected values, with wider ranges of results from spatial error models of primary pollutants (Figure A.3). The range of simulation results is consistent with the observed variability in yearly inter-monitor covariance.

3.3.5 Error Impact on Health Risk Assessment

For the base case of 1-hr maximum CO exposures (Z^*) and CVD outcomes, a risk ratio of 1.0139 was observed, with p-value of 0.000009. Differences between this result and results from the regression of the same health outcome data against the simulations with error added instead of the CO data reflect the impact of error. Results of 1000 epidemiologic models for 12 air pollutants and three error scenarios are summarized in Table 2. The degree to which the risk ratio observed using the simulations is decreased toward one (the null hypothesis) from the risk ratio observed using the base case data can be expressed as a percent bias-to-null in the risk ratio.

$$RR\ percent bias\text{-}to\text{-}null = \left(\frac{e^{\beta^*} - e^{\beta}}{e^{\beta^*} - 1} \right) \times 100\% \approx \left(1 - \frac{\beta}{\beta^*} \right) \times 100\% \qquad (3.14)$$

Here, β^* is the effect estimator when CO central monitor data (Z^*) were used in the epidemiologic model and β is the effect estimator when simulations with error added (Z) were used.

Table 3.2. Average risk ratios and median p-values for 1000 trials with error added to base case. Base case results: RR = 1.0139, p-value = 0.000009. For spatial error, interquartile ranges are shown in parentheses.

	instrument error		spatial error		spatial error (no autocorrelation)	
	RR per ppm	p-value	RR per ppm	p-value	RR per ppm	p-value
1-hr max NO_2	1.0133	0.000021	1.0046 (1.0026 - 1.0065)	0.1038 (0.0236 - 0.362)	1.0044	0.1196
1-hr max NO_x	1.0132	0.000018	1.0079 (1.0057 - 1.0100)	0.0263 (0.0053 - 0.119)	1.0074	0.0321
8-hr max O_3	1.0139	0.000010	1.0128 (1.0118 - 1.0139)	0.000050 (0.000011 - 0.00018)	1.0126	0.000059
1-hr max SO_2	1.0132	0.000026	1.0045 (1.0023 - 1.0065)	0.1411 (0.0355 - 0.425)	1.0042	0.1695
1-hr max CO	1.0131	0.000020	1.0059 (1.0039 - 1.0077)	0.0392 (0.0073 - 0.172)	1.0049	0.0465
24-hr PM_{10}	1.0131	0.000027	1.0103 (1.0085 - 1.012)	0.0017 (0.00025 - 0.0089)	1.0099	0.0021
24-hr $PM_{2.5}$	1.0138	0.000016	1.0126 (1.0113 - 1.0139)	0.00015 (0.000027 - 0.00070)	1.0123	0.00019
24-hr $PM_{2.5}$-SO_4	1.0134	0.000015	1.0121 (1.0109 - 1.0133)	0.00011 (0.000019 - 0.00047)	1.0118	0.00012
24-hr $PM_{2.5}$-NO_3	1.0143	0.000011	1.0145 (1.0128 - 1.016)	0.00015 (0.000032 - 0.00081)	1.0141	0.00020
24-hr $PM_{2.5}$-NH_4	1.0141	0.000020	1.0117 (1.0101 - 1.0134)	0.00057 (0.000091 - 0.0032)	1.0113	0.00076
24-hr $PM_{2.5}$-EC	1.0134	0.000022	1.0079 (1.0050 - 1.0107)	0.0440 (0.0071 - 0.192)	1.0078	0.0420
24-hr $PM_{2.5}$-OC	1.0132	0.000045	1.0101 (1.0084 - 1.0118)	0.0017 (0.00024 - 0.0092)	1.0097	0.0021

The addition of instrument precision error to monitor data was found to have little impact on risk ratio and significance estimates. Average risk ratios ranged from 1.0131 to 1.0143 (compared to 1.0139), and median p-values ranged from 0.000010 to 0.000045 (compared to 0.000009). In comparison, the impact of error due to spatial variability was much larger. Loss of significance of association at the 95% level was observed for SO_2 and NO_2. For primary pollutants, risk ratios on average ranged from 1.0045 for SO_2 to 1.0079 for NO_x, corresponding to average bias-to-null estimates of 68 to 43%, respectively. For pollutants largely of secondary origin, the results exhibited less bias-to-null, with risk ratios on average ranging from 1.0117 for NH_4 to 1.0145 for NO_3; these correspond to average bias-to-null estimates of less than 16%. For pollutants of mixed

origin, intermediate levels of bias-to-null were observed. The presence of temporal autocorrelation in spatial error slightly decreased the bias-to-null and increased significance, likely due to the fact that the air pollution data themselves have temporal autocorrelation.

In Figure 3.4, average values of the bias-to-null from all models are shown as a function of the integrated scaled semivariance, which is a measure of the magnitude of error added. These results support our hypothesis that the impact of error due to spatial variability depends on how large the spatial variance is relative to the temporal variance, evaluated for the entire study population over the entire study period.

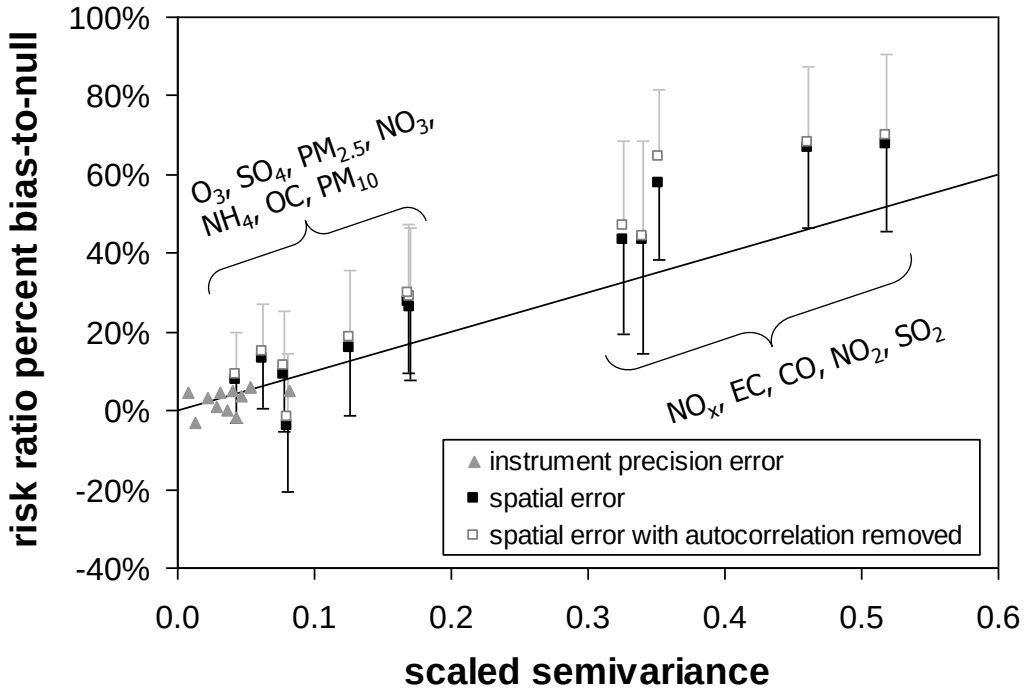


Figure 3.4. Percent reduction in risk ratio due to instrument precision error and spatial variability error versus semivariogram nugget (γ_o) and integrated population-weighted semivariance ($\overline{\gamma}'$), respectively, on concentration basis with one-sided error bars indicating the standard deviation of the 1000 simulations. Spatial variability error points are labeled in order of increasing $\overline{\gamma}'$. For reference, a one-to-one line is shown.

This investigation was limited to error associated with the lack of correlation between exposure measurements. Results are limited to single pollutant analyses. Our results can be used directly to re-interpret results from single pollutant models across multiple pollutants, with health effect estimator attenuation varying across pollutants. Our modeling framework could be extended for use in multi-pollutant models as well by including joint variance effects in the semivariogram analysis.

In this paper, we have shown the relative impacts of measurement error on health risk assessment using simulated time-series of ambient levels. The amount of error was derived from observations; however, bias in the health response estimate is dependent on both the amount and type of measurement error. In terms of the latter, two extremes in the conceptual framework for error type have been proposed: classical error and Berkson error (Sheppard et al. 2005; Zeger et al. 2000). The type of error simulated in this work is neither classical nor Berkson (Figures A.4 and A.5). In ongoing work, we are exploring the impact of varying error type, while keeping the amount of error constant, on health effect estimation by relaxing the assumption that the simulations with error added have the same distribution as the true values.

CHAPTER 4

IMPACT OF EXPOSURE MEASUREMENT ERROR IN AIR POLLUTION EPIDEMIOLOGY: EFFECT OF ERROR TYPE IN TIME-SERIES STUDIES

(Goldman GT, Mulholland JA, Russell AG, Strickland MJ, Klein M, Waller LA, Tolbert PE. *Environmental Health*, 10: 61)

Abstract

Background

Two distinctly different types of measurement error are Berkson and classical. Impacts of measurement error in epidemiologic studies of ambient air pollution are expected to depend on error type. We characterize measurement error due to instrument imprecision and spatial variability as multiplicative (i.e. additive on the log scale) and model it over a range of error types to assess impacts on risk ratio estimates both on a per measurement unit basis and on a per interquartile range (IQR) basis in a time-series study in Atlanta.

Methods

Daily measures of twelve ambient air pollutants were analyzed: NO_2, NO_x, O_3, SO_2, CO, PM_{10} mass, $PM_{2.5}$ mass, and $PM_{2.5}$ components sulfate, nitrate, ammonium, elemental carbon and organic carbon. Semivariogram analysis was applied to assess spatial variability. Error due to this spatial variability was added to a reference pollutant time-series on the log scale using Monte Carlo simulations. Each of these time-series was exponentiated and introduced to a Poisson generalized linear model of cardiovascular disease emergency department visits.

Results

Measurement error resulted in reduced statistical significance for the risk ratio estimates for all amounts (corresponding to different pollutants) and types of error. When modeled as classical-type error, risk ratios were attenuated, particularly for primary air pollutants, with average attenuation in risk ratios on a per unit of measurement basis ranging from 18% to 92% and on an IQR basis ranging from 18% to 86%. When modeled as Berkson-type error, risk ratios per unit of measurement were biased away from the null hypothesis by 2% to 31%, whereas risk ratios per IQR were attenuated (i.e. biased toward the null) by 5% to 34%. For CO modeled error amount, a range of error types were simulated and effects on risk ratio bias and significance were observed.

Conclusions

For multiplicative error, both the amount and type of measurement error impact health effect estimates in air pollution epidemiology. By modeling instrument imprecision and spatial variability as different error types, we estimate direction and magnitude of the effects of error over a range of error types.

4.1 Background

The issue of measurement error is unavoidable in epidemiologic studies of air pollution (Sarnat et al. 2007). Although methods for dealing with this measurement error have been proposed (Carroll et al. 1995; Fuller 1987) and applied to air pollution epidemiology specifically (Dominici et al. 2000; Strand et al. 2006), the issue remains a central concern in the field (Ren and Tong 2008). Because large-scale time-series studies often use single central monitoring sites to characterize community exposure to ambient concentrations (Wilson et al. 2005), uncertainties arise regarding the extent to which these monitors are representative of exposure. Zeger et al. (Zeger et al. 2000) identify three components of measurement error: (1) the difference between individual exposures and average personal exposure, (2) the difference between average personal exposure and

ambient levels, and (3) the difference between measured and true ambient concentrations. While the former two components of error can have a sizeable impact on epidemiologic findings that address etiologic questions of health effects and personal exposure, it is the third component that is particularly relevant in time-series studies that address questions of the health benefits of ambient regulation (Carrothers and Evans 2000).

Prior studies have suggested that the impact of measurement error on time-series health studies differs depending upon the type of error introduced (Armstrong 1998; Sheppard et al. 2005; Zeger et al. 2000). Two distinctly different types of error have been identified. One type is classical error, in which measurements, Z_t, vary randomly about true concentrations, Z_t^*; this can be considered the case for instrument error associated with ambient monitors. That is, instrument error is independent of the true ambient level, such that $E[Z_t \mid Z_t^*] = Z_t^*$. Moreover, the variation in the measurements, Z_t, is expected to be greater than the variation in the true values, Z_t^*. Therefore, classical error is expected to attenuate the effect estimate in time-series epidemiologic studies. In contrast, under a Berkson error framework, the true ambient, Z_t^*, varies randomly about the measurement, Z_t. This might be the case, for example, of a measured population average over the study area with true individual ambient levels varying randomly about this population average measurement. In this case, measurement error is independent of the measured population average ambient; that is, $E[Z_t^* \mid Z_t] = Z_t$. Furthermore, the measurement, Z_t, is less variable than the true ambient level, Z_t^*. A purely Berkson error is expected to yield an unbiased effect estimate, provided that the true dose-response is linear (Fuller 1987).

Several studies have investigated the impact of error type on regression models. The simultaneous impact of classical and Berkson errors in a parametric regression estimating radon exposure has been investigated (Reeves et al. 1998) and error type has

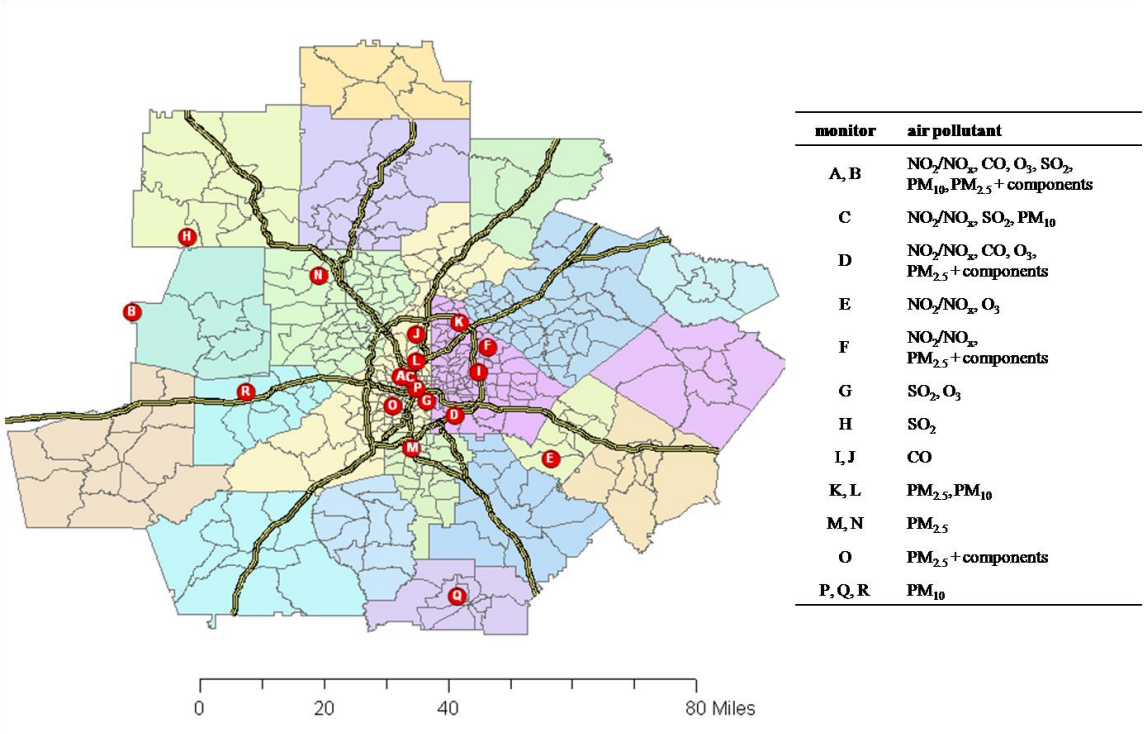
been assessed in a semiparametric Bayesian setting looking at exposure to radiation from nuclear testing (Li et al. 2007; Mallick et al. 2002); however, no study to date has comprehensively assessed the impact of error type across multiple pollutants for instrument imprecision and spatial variability in a time-series context.

Error type depends on the relationship between the distribution of measurements and the distribution of true values. Because true relevant exposure in environmental epidemiologic studies is not known exactly, determination of error type is challenging; thus, here we examine the impact of error modeled as two distinctly different types: classical and Berkson. First, we examine monitor data to assess whether error is better modeled on a logged or unlogged basis. Typically, researchers investigating error type have added error on an unlogged basis, e.g. (Armstrong 1998; Zeger et al. 2000); however, air pollution data are more often lognormal due to atmospheric dynamics and concentration levels that are never less than zero. It is plausible that true ambient exposures are distributed lognormally about a population average as well; therefore, measurement error may be best described as additive error on the log scale. We investigate the combined error from two sources that have been previously identified as relevant in time-series studies: (1) instrument precision error and (2) error due to spatial variability (Carrothers and Evans 2000). We limit our scope to ambient levels of pollutants measured in accordance with regulatory specifications, disregarding spatial microscale variability, such as near roadway concentrations, as well as temporal microscale variability, such as that associated with meteorological events on sub-hour time scales. Here, building on a previously developed model for the amount of error associated with selected ambient air pollutants (Goldman et al. 2010), we quantitatively assess the effect of error type on the impacts of measurement error on epidemiologic results from an ongoing study of air pollution and emergency department visits in Atlanta.

4.2 Methods

4.2.1 Air Pollutant Data

Daily metrics of 12 ambient air pollutants were studied: 1-hr maximum NO_2, NO_x, SO_2 and CO, 8-hr maximum O_3, and 24-hr average PM_{10}, $PM_{2.5}$ and $PM_{2.5}$ components sulfate (SO_4), nitrate (NO_3),ammonium (NH_4),elemental carbon (EC) and organic carbon (OC). Observations were obtained from three monitoring networks: the US EPA's Air Quality System (AQS), including State and Local Air Monitoring System and Speciation Trends Network for $PM_{2.5}$ component measurements; the Southeastern Aerosol Research and Characterization Study (SEARCH) network (Hansen et al. 2003), including the Atlanta EPA supersite at Jefferson Street (Solomon et al. 2003b)**;** and the Assessment of Spatial Aerosol Composition in Atlanta (ASACA) network (Butler et al. 2003). Locations of the monitoring sites are shown in Figure 4.1.



monitor	air pollutant
A, B	NO_2/NO_x, CO, O_3, SO_2, PM_{10}, $PM_{2.5}$ + components
C	NO_2/NO_x, SO_2, PM_{10}
D	NO_2/NO_x, CO, O_3, $PM_{2.5}$ + components
E	NO_2/NO_x, O_3
F	NO_2/NO_x, $PM_{2.5}$ + components
G	SO_2, O_3
H	SO_2
I, J	CO
K, L	$PM_{2.5}$, PM_{10}
M, N	$PM_{2.5}$
O	$PM_{2.5}$ + components
P, Q, R	PM_{10}

Figure 4.1. Map of 20-county metropolitan Atlanta study area. Census tracts, expressways, and ambient air pollutant monitoring sites are shown.

To assess error due to instrument imprecision and spatial variability of ambient concentrations, 1999-2004 datasets were used for the 12 pollutants with data completeness for this time period (2,192 days) ranging from 82% to 97%. Data from collocated instruments were used to characterize instrument precision error. Measurement methods and data quality are discussed in detail in our prior work (Goldman et al. 2010). Distributions of all air pollutant measures more closely approximate lognormal distributions than normal distributions (see Table B.1); therefore, additive error was characterized and modeled on a log concentration basis so that simulations with error added to a base case time-series would also have lognormal distributions.

4.2.2 Measurement Error Model

The measurement error model description here highlights differences from our previous work in which error type effects were not addressed (Goldman et al. 2010). In this study, a time-series of observed data was taken to be the "true" time-series, Z_t^*, serving as a base case. Classical-like or Berkson-like error was added to this base case to produce a simulated time-series, Z_t, that represents a population-weighted average ambient time-series. Here, the asterisk refers to a true value (i.e. without error) as opposed to a value that contains error (i.e. the simulated values in this study). The choice of which pollutant to use for the true, or base case, time-series is arbitrary, as long as an association with a health endpoint has been observed with that pollutant. To develop simulated datasets with modeled instrument and spatial error added, the following steps were taken. Base case time-series data were normalized as follows.

$$\chi_t^* = \frac{\ln Z_t^* - \mu_{\ln Z^*}}{\sigma_{\ln Z^*}} \tag{4.1}$$

Here, χ_t^* is the normalized log concentration on day *t* and $\mu_{\ln Z^*}$ and $\sigma_{\ln Z^*}$ are the mean and standard deviation, respectively, of the log concentrations over all days *t*; thus, the

mean and standard deviation of χ_t^* are 0 and 1, respectively. Error in χ_t^* was modeled as multiplicative (i.e. additive on a log scale) as follows.

$$\varepsilon_{\chi t} = N_t \sigma_{err} \tag{4.2}$$

Here, $\varepsilon_{\chi t}$ is the modeled error in χ_t^* for day *t*, N_t is a random number with distribution $\sim N(0,1)$ and σ_{err} is the standard deviation of error added, a parameter derived from the population-weighted semivariance to capture the amount of error present for each pollutant, as described in the next subsection. Short-term temporal autocorrelation observed in the differences between measurements was modeled using a three-day running average of random numbers for N_t (Goldman et al. 2010).

To provide simulations of monitor data with error added (Z_t), the modeled error was added to normalized data and then the normalized data with error added were denormalized in two ways: one to simulate classical-like error (i.e. classical error on a log concentration basis, referred to here as type C error) and the other to simulate Berkson-like error (i.e. Berkson error on a log concentration basis, referred to here as type B error). Simulations with type C error are generated by eq. 4.3.

$$\text{type C error: } \chi_t = \chi_t^* + \varepsilon_{\chi t} \tag{4.3}$$

Here, χ_t is the standardized simulated time-series (on the log scale) with type C error added and normal distribution $\sim N\left(0, \sqrt{1+\sigma_{err}{}^2}\right)$. In this case of type C error, $\varepsilon_{\chi t}$ and χ_t^* are independent (i.e. $E[R(\varepsilon_{\chi t}, \chi_t^*)] = 0$). For type B error, $\varepsilon_{\chi t}$ and χ_t are independent (i.e. $E[R(\varepsilon_{\chi t}, \chi_t)] = 0$) and $\chi_t^* = \chi_t + \varepsilon_{\chi t}$. It can be shown (see eqs. B.1-B.6) that simulations with type B error can be generated from the true time-series by eq. 4.4.

$$\text{type B error: } \chi_t = (\chi_t^* + \varepsilon_{\chi t})/(1+\sigma_{err}{}^2) \tag{4.4}$$

Here, χ_t is the standardized simulated time-series (on the log scale) with type B error added and normal distribution $\sim N\left(0, \frac{1}{\sqrt{1+\sigma_{err}{}^2}}\right)$. After the standardized simulated time-series is generated by either eq. 4.3 or eq. 4.4, the simulations are denormalized by eq. 4.5.

$$Z_t = \exp(\chi_t \sigma_{\ln Z^*} + \mu_{\ln Z^*}) \tag{4.5}$$

For both error types, the simulated time-series (Z_t) and true time-series (Z_t^*) have the same log means ($\mu_{\ln Z} = \mu_{\ln Z^*}$). For classical-like error (type C), the log standard deviation is greater for the simulated time-series than the true time-series ($\sigma_{\ln Z} > \sigma_{\ln Z^*}$) because the simulated values are scattered about the true values. For Berkson-like error (type B), the log standard deviation is less for the simulated time-series than the true time-series ($\sigma_{\ln Z} < \sigma_{\ln Z^*}$) because the true values are scattered about the simulated values.

4.2.3 Semivariogram Analysis

To quantify the amount of error (i.e. σ_{err}) due to instrument imprecision and spatial variability to add to the simulated time-series for each pollutant (eq. 4.2), we made use of the geostatistical tool of the semivariogram, which provides information on spatial autocorrelation of data and has proved useful in air pollution applications (Casado et al. 1994; Wade et al. 2006). Here, the semivariance of the differences between normalized observations (χ_k and χ_l) at two locations (*k* and *l*) located a distance *h* apart is normalized by the temporal variance (variation over the time-series of observations) of the average of two normalized observations to yield a scaled semivariance, γ'. It can be shown that this scaled semivariance (i.e. the semivariance of normalized values) is related

to the Pearson correlation coefficient (R) between normalized observations from two monitors as follows (Wade et al. 2006).

$$\gamma'(h) \equiv \frac{Var\left(\frac{\chi_k - \chi_l}{2}\right)}{Var\left(\frac{\chi_k + \chi_l}{2}\right)} = \frac{1 - R(h)}{1 + R(h)} \tag{4.6}$$

Thus, γ' represents the spatial semivariance scaled to a quantity indicative of the range of exposures over which health risk is being assessed; it is unitless and allows for comparison across pollutants. A scaled semivariance value of 0 corresponds to perfectly correlated observations (R = 1) and a value of 1 corresponds to perfectly uncorrelated observations (R = 0).

Correlations between observations from all pairs of monitors measuring the same pollutant during 1999-2004 were calculated on a log concentration basis. Assuming the spatial variation of air pollutants to be isotropic, scaled semivariograms were constructed and modeled as a function of the distance between observations, *h*, using a sill of 1, nugget values derived from collocated measurement time-series described in previous work, and least squares regression to determine the range (Goldman et al. 2010). The estimate from the semivariogram function for each of the 660 Census tracts was weighted by the population in that tract (estimates from 2000 Census data) to derive an overall population-weighted average for each pollutant; thus, the population-weighted semivariance includes impacts of both instrument imprecision and spatial variability and represents the population-weighted average semivariance between all residences in the study area.

$$\overline{\gamma'} = \frac{1}{p_{total}}\left(\sum_{i}^{660}\sum_{i+1}^{660} p_{i,j}\gamma'_{i,j} + \sum_{i}^{660} p_{i,i}\gamma'_{i,i}\right) \tag{4.7}$$

Here, $\overline{\gamma'}$ is the population-weighted average scaled semivariance on a log scale, p_{total} is the total population of the study area, $p_{i,j}$ is the sum of population in census tracts *i* and *j*,

and $\gamma'_{i,j}$ is the value of the semivariance function at the distance between centroids of census tracts *i* and *j*. For within-tract resident pairs, an average distance between residences was applied. Semivariograms for each of the twelve pollutants studied have been shown previously (Goldman et al. 2010) and population-weighted semivariances are in Table 4.1. The population-weighted semivariance is related to the population-weighted correlation coefficient as follows.

$$\overline{\gamma'} = \frac{1-\overline{R}}{1+\overline{R}} \tag{4.8}$$

Table 4.1. Population-weighted scaled semivariances, $\overline{\gamma'}$, Pearson correlation coefficients, $\overline{R}$, and model parameters used in the Monte Carlo simulations to simulate amount of error (σ_{err}) and error type ($\sigma_{\ln Z}/\sigma_{\ln Z^*}$).

Pollutant	$\overline{\gamma'}$	$\overline{R}$	σ_{err}	$\sigma_{\ln Z}/\sigma_{\ln Z^*}$ Type B	$\sigma_{\ln Z}/\sigma_{\ln Z^*}$ Type C
1-hr max NO_2	0.516	0.320	1.46	0.57	1.77
1-hr max NO_x	0.384	0.445	1.12	0.67	1.50
8-hr max O_3	0.051	0.903	0.33	0.95	1.05
1-hr max SO_2	0.517	0.319	1.46	0.56	1.77
1-hr max CO	0.411	0.418	1.18	0.65	1.55
24-hr PM_{10}	0.192	0.678	0.69	0.82	1.21
24-hr $PM_{2.5}$	0.100	0.819	0.47	0.90	1.11
24-hr $PM_{2.5}$-SO_4	0.068	0.873	0.38	0.93	1.07
24-hr $PM_{2.5}$-NO_3	0.140	0.754	0.57	0.87	1.15
24-hr $PM_{2.5}$-NH_4	0.149	0.741	0.59	0.86	1.16
24-hr $PM_{2.5}$-EC	0.337	0.495	1.01	0.70	1.42
24-hr $PM_{2.5}$-OC	0.175	0.702	0.65	0.84	1.19

Model parameter σ_{err} (eq. 4.2) is defined to provide simulations with an amount of error such that $E[R(\ln Z, \ln Z^*)] = \sqrt{\overline{R}}$ where $\overline{R}$ is obtained from semivariogram analysis (eqs. 4.6-4.8). The correlation between the true ambient time-series and a time-series with error added, i.e. R(ln *Z*, ln *Z**), is the square root of the correlation between any two time-series, i.e. R(ln Z_1, ln Z_2), where each is derived by adding the same amount of error to the true ambient time-series. Since the standard deviation of χ_t depends on σ_{err}, the standard deviation of the simulated time-series relative to that of the true time-

series ($\sigma_{\ln Z}/\sigma_{\ln Z^*}$) depends on $\overline{R}$ as well. The following analytical relationships for σ_{err} and $\sigma_{\ln Z}/\sigma_{\ln Z^*}$ were derived (see eqs. B.7-B.10).

$$\sigma_{err} = \sqrt{\frac{2\overline{\gamma'}}{1-\overline{\gamma'}}} = \sqrt{\frac{1-\overline{R}}{\overline{R}}} \tag{4.9}$$

$$\frac{\sigma_{\ln Z}}{\sigma_{\ln Z^*}} = \begin{cases} \sqrt{\frac{1+\overline{\gamma'}}{1-\overline{\gamma'}}} = \frac{1}{\sqrt{\overline{R}}} > 1 & \text{for error type C} \\ \sqrt{\frac{1-\overline{\gamma'}}{1+\overline{\gamma'}}} = \sqrt{\overline{R}} < 1 & \text{for error type B} \end{cases} \tag{4.10}$$

Values of σ_{err} and $\sigma_{\ln Z}/\sigma_{\ln Z^*}$ used here can be found in Table 4.1.

Sets of 1000 simulated time-series with instrument and spatial error added for each pollutant for the scenarios of C and B error types were produced for the six-year period 1999-2004. In addition, simulations of CO measurement error only were generated for a range of error types with $\sigma_{\ln Z}/\sigma_{\ln Z^*}$ values between error types C and B. In Figure 4.2, examples of scatterplots demonstrate that C and B error types defined on a log basis (i.e. $\ln Z - \ln Z^*$) are independent of $\ln Z^*$ and $\ln Z$, respectively.

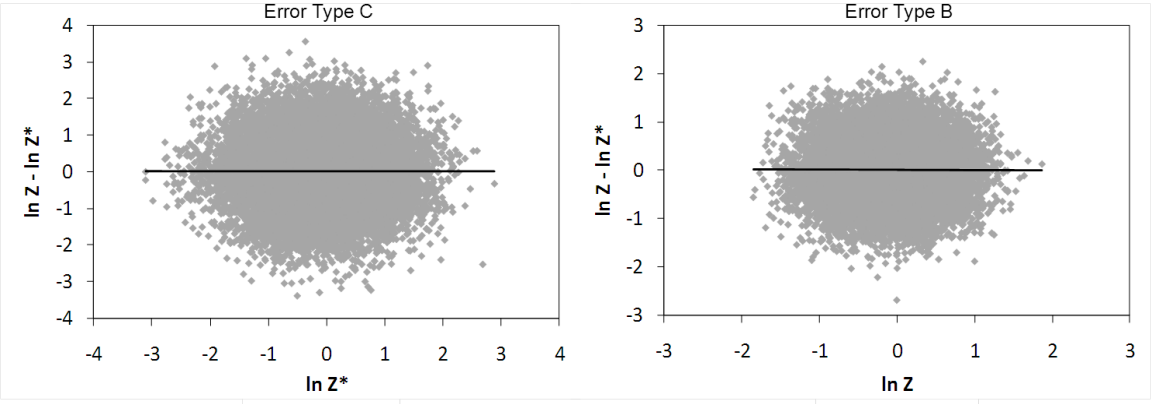


Figure 4.2. Scatterplots of CO error ($\overline{\gamma'} = 0.411$) versus $\ln Z^*$ for error type C (left panel) and versus $\ln Z$ for error type B (right panel). Black lines indicate lines of linear regression.

4.2.4 Epidemiologic Model

Relationships between daily measures of ambient air pollution and daily counts of emergency department (ED) visits for cardiovascular disease (CVD, including ischemic

heart disease, dysrhythmia, congestive heart failure, and peripheral/cerebrovascular disease) were assessed using methods described elsewhere (Metzger et al. 2004) and briefly summarized here. There were 166,950 ED visits for CVD in the 20-county metropolitan Atlanta area during 1999-2004. Lag 0 associations between daily pollutant concentration and the daily count of ED visits were assessed using Poisson generalized linear models that were scaled to accounted for overdispersion. The general form of the epidemiologic model is

$$\log[\mathrm{E}(Y_t)] = \alpha + \beta\ Z_t + \gamma\, confounders_t \tag{4.11}$$

where Y_t is the count of emergency department visits, Z_t is the mismeasured pollutant concentration, and $confounders_t$ is the vector of potential confounders on day t. The specific potential confounders included in the model were indicator variables for day-of-week, season, and when a hospital entered or left the study; cubic terms for maximum temperature and dew point; and a cubic spline with monthly knots for day of follow-up. Poisson regression yields α as the intercept, β as the log of the rate ratio associated with a unit change in pollutant concentration, and γ as the vector of regression coefficients for the suspected confounders included in the model. The risk ratios (RR) per unit of measurement change and per interquartile range (IQR) change in pollutant concentration (Z) are given by eq. 4.12 and eq. 4.13, respectively.

$$RR\ per\ unit = e^{\beta} \tag{4.12}$$

$$RR \text{ per IQR } = e^{IQR\times\beta} \tag{4.13}$$

Using data from the central monitor, preliminary epidemiologic assessments were performed for all air pollutants and ED visits for CVD. Consistent with previous findings (Metzger et al. 2004), significant positive associations were found for several traffic-related pollutants, including NO_x, CO and EC. For the measurement error analysis described here, we used 1-hr maximum CO data as our base case, representing in our analysis a true time-series and the measured risk ratio the true association. In this way,

the exposure and health outcome values that we chose to represent true time-series have distributional characteristics expected of ambient air pollution and ED visit data. Simulations with measurement error added to the base case were used to evaluate the impact of measurement error on the epidemiologic analyses. A Monte Carlo approach was used to assess uncertainty. As already described, the relationship between this base case time-series and a simulated time-series is that expected of the average relationship between the true ambient time-series for all people and a population-weighted average time-series based on measurements in terms of error amount, with different error types evaluated. A percent attenuation in risk ratio (toward the null hypothesis of 1) is calculated as follows, with RR^* representing the true risk ratio (obtained from the base case Poisson regression) and RR representing the risk ratio obtained using simulated population-weighted time-series.

$$\text{percent attenuation in RR} = \left(\frac{RR^* - RR}{RR^* - 1}\right) \times 100\% \qquad (4.14)$$

4.3 Results

4.3.1 Distribution of Measurement Error Simulations

Analysis of the distributions of correlation coefficients between the true log concentrations (i.e. the base case) and the simulated log concentrations, $R(\ln Z, \ln Z^*)$, and between the log error, $\varepsilon_{\ln Z} = \ln Z - \ln Z^*$, and the true and simulated log concentrations, $R(\varepsilon_{\ln Z}, \ln Z^*)$ and $R(\varepsilon_{\ln Z}, \ln Z)$, for 1000 simulations for each pollutant and each error type demonstrates that the simulations contain on average the desired amounts and types of error (Figures 4.3 and 4.4). Wider distributions were observed for more spatially heterogeneous pollutants.

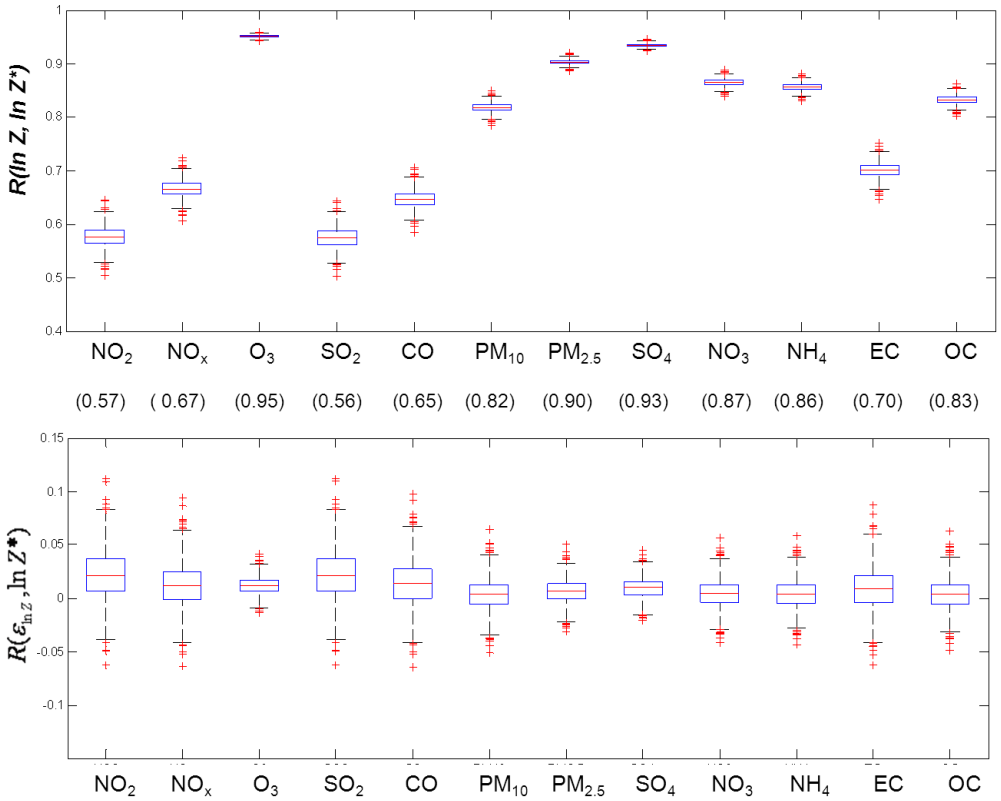


Figure 4.3. Boxplots of $R(\ln Z, \ln Z^*)$ (top panel), with expected correlation coefficients shown in parentheses, and $R(\varepsilon_{\ln Z}, \ln Z^*)$ (bottom panel) for 1000 simulated data time-series of error type C simulations.

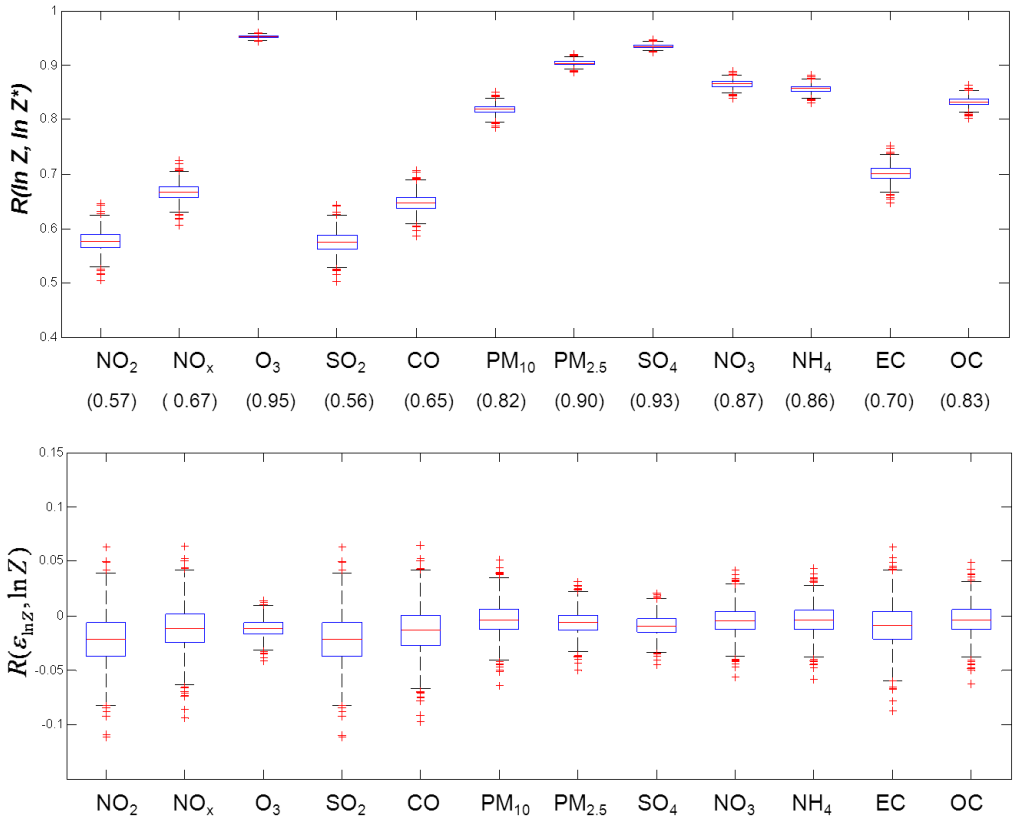


Figure 4.4. Boxplots of $R(\ln Z, \ln Z^*)$ (top panel), with expected correlation coefficients shown in parentheses, and $R(\varepsilon_{\ln Z}, \ln Z)$ (bottom panel) for 1000 simulated data time-series of error type B simulations.

4.3.2 Impact of Error on Health Assessment

For the base case of 1-hour maximum CO measurements and CVD outcomes, a RR per ppm of 1.0139 was observed, with a 95% confidence interval (CI) of 1.0078-1.0201 and a p-value of 0.000009. With an IQR of 1.00 ppm, the RR per IQR and corresponding CI are the same as those on a per unit of measurement basis for our base case. For epidemiologic models using the time-series with simulated error added, the RR and CI results are not the same on a per measurement unit basis and a per IQR basis because the IQR of the simulated values is not 1. As expected, the simulated time-series with error type C has a greater IQR than the base case since this error is scattered about the true values, and the simulated time-series with error type B has a lower IQR than the

base case since this error is scattered about the simulated values. Results of 1000 epidemiologic models for each of 12 air pollutants and two error scenarios are summarized in Table 4.2. The reported p-values represent those calculated from average z-score statistics and 95% confidence intervals were calculated using the asymptotic standard error estimates obtained from the regression model.

Table 4.2. Summarized epidemiologic model results with the magnitude of error representative of error associated with using a population-weighted average for each pollutant added to the base case (RR^* = 1.0139, 95% CI = 1.0078-1.0201, p-value = 0.000009, IQR = 1.00 ppm).

pollutant	RR per ppm (95% CI)	IQR (ppm)	RR per IQR (95% CI)	p-value
Error Type C simulations				
1-hr max NO_2	1.0011 (0.9998-1.0023)	1.84	1.0020 (0.9997-1.0042)	0.0957
1-hr max NO_x	1.0024 (1.0003-1.0046)	1.51	1.0037 (1.0005-1.0070)	0.0251
8-hr max O_3	1.0114 (1.0060-1.0169)	1.05	1.0120 (1.0063-1.0178)	0.00004
1-hr max SO_2	1.0011 (0.9998-1.0023)	1.84	1.0019 (0.9997-1.0042)	0.0966
1-hr max CO	1.0021 (1.0002-1.0040)	1.57	1.0033 (1.0003-1.0063)	0.0342
24-hr PM_{10}	1.0063 (1.0025-1.0102)	1.20	1.0076 (1.0030-1.0122)	0.0013
24-hr $PM_{2.5}$	1.0094 (1.0045-1.0142)	1.10	1.0103 (1.0049-1.0156)	0.000157
24-hr $PM_{2.5}$-SO_4	1.0107 (1.0054-1.0159)	1.07	1.0114 (1.0058-1.0170)	0.000066
24-hr $PM_{2.5}$-NO_3	1.0079 (1.0035-1.0123)	1.14	1.0090 (1.0040-1.0141)	0.00040
24-hr $PM_{2.5}$-NH_4	1.0076 (1.0033-1.0119)	1.15	1.0088 (1.0038-1.0137)	0.00050
24-hr $PM_{2.5}$-EC	1.0032 (1.0006-1.0057)	1.42	1.0045 (1.0009-1.0081)	0.0140
24-hr $PM_{2.5}$-OC	1.0068 (1.0028-1.0108)	1.18	1.0080 (1.0033-1.0128)	0.00090
Error Type B simulations				
1-hr max NO_2	1.0182 (1.0041-1.0325)	0.51	1.0092 (1.0021-1.0165)	0.0112
1-hr max NO_x	1.0169 (1.0056-1.0284)	0.61	1.0103 (1.0034-1.0172)	0.0034
8-hr max O_3	1.0142 (1.0075-1.0208)	0.94	1.0133 (1.0070-1.0195)	0.000027
1-hr max SO_2	1.0182 (1.0041-1.0325)	0.51	1.0092 (1.0021-1.0164)	0.0114
1-hr max CO	1.0172 (1.0053-1.0292)	0.59	1.0101 (1.0031-1.0171)	0.0044
24-hr PM_{10}	1.0152 (1.0068-1.0236)	0.78	1.0117 (1.0053-1.0182)	0.00030
24-hr $PM_{2.5}$	1.0144 (1.0073-1.0217)	0.88	1.0127 (1.0064-1.0190)	0.000074
24-hr $PM_{2.5}$-SO_4	1.0143 (1.0074-1.0211)	0.92	1.0130 (1.0068-1.0193)	0.000039
24-hr $PM_{2.5}$-NO_3	1.0147 (1.0071-1.0225)	0.83	1.0122 (1.0059-1.0186)	0.000152
24-hr $PM_{2.5}$-NH_4	1.0148 (1.0070-1.0226)	0.82	1.0121 (1.0058-1.0185)	0.000175
24-hr $PM_{2.5}$-EC	1.0165 (1.0060-1.0271)	0.65	1.0106 (1.0038-1.0174)	0.0021
24-hr $PM_{2.5}$-OC	1.0150 (1.0069-1.0232)	0.79	1.0119 (1.0055-1.0183)	0.00030

When instrument imprecision and spatial variability error were added as error type C, the average IQR of simulated time-series was greater than the IQR of the base case for all pollutants; for error type B, the average IQR of simulated time-series was less than the IQR of the base case for all pollutants. As expected, adding error to the base case resulted in a reduction of significance (i.e. a higher p-value) for both error types, as

shown graphically in Figure 4.5. The greater the amount of error (i.e. the greater the population-weighted semivariance), the greater the reduction in significance observed. Primary pollutants (SO_2, NO_2/NO_x, CO, and EC) had more error than secondary pollutants and those of mixed origin (O_3, SO_4, NO_3, NH_4, $PM_{2.5}$, OC, and PM_{10}) due to greater spatial variability. Regarding error type, there was a greater reduction of statistical significance when error type was modeled as type C than when error type was modeled as type B. For NO_2 and SO_2, which have the largest amount of measurement error, there was a loss of significance (p-value > 0.05) when error was modeled as error type C.

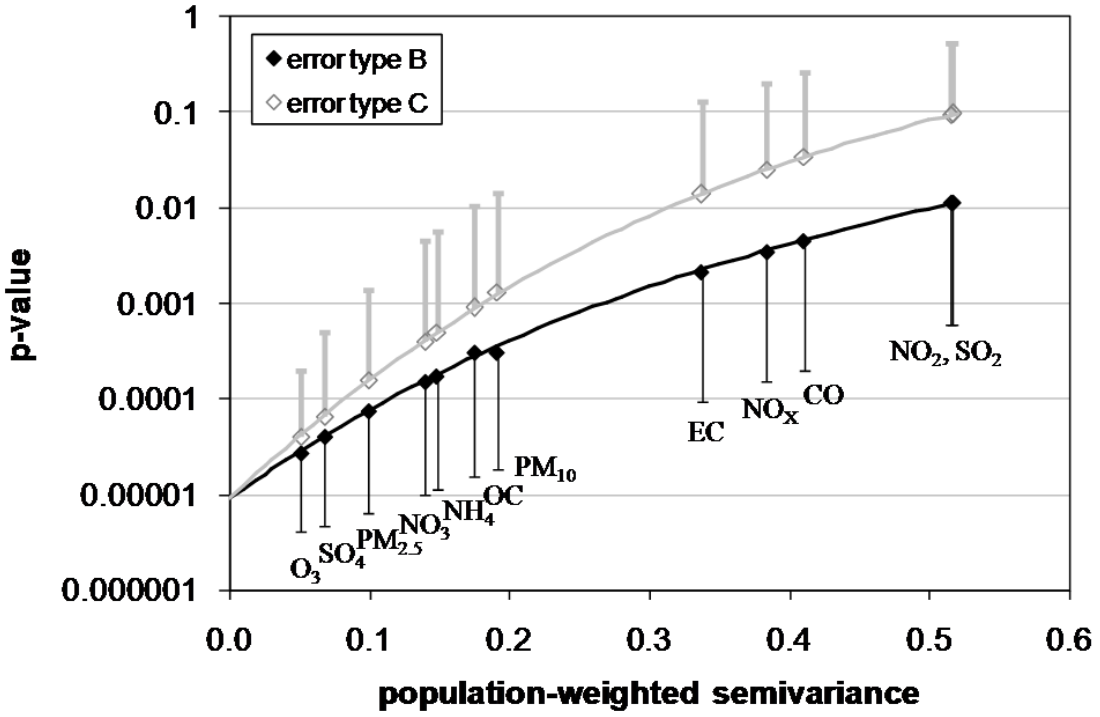


Figure 4.5. P-values versus population-weighted semivariance. Half-bars denote standard deviations for 1000 error simulations.

Risk ratio results for the two error types are plotted in Figure 4.6 on a percent attenuation basis. RR per unit of measurement decreased, and attenuation increased, with increasing error added (i.e. increasing population-weighted semivariance) when the error was of type C. However, RR per unit increased, with increasing bias away from the null, with increasing error added when error was of type B. For NO_2 and SO_2, which had the

most measurement error, the attenuation was 92% when modeled as error type C and biased away from the null by 31% when modeled as error type B. On a per IQR basis, variation in the RR estimates between error types was much less dramatic. Both error types C and B led to lower RR estimates (i.e. bias towards the null). For NO_2 and SO_2, which again had the most measurement error, the attenuation was 86% when modeled as type C and 34% when modeled as type B error. For error type B there was a wider distribution of results than for type C error.

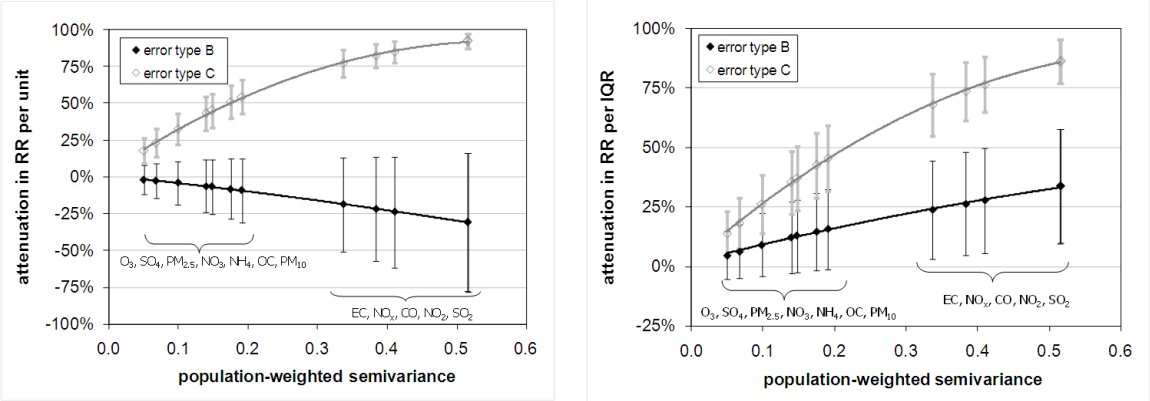


Figure 4.6. Percent attenuation in risk ratio per ppm (left panel) and per IQR (right panel) due to error versus population-weighted semivariance. Bars denote standard deviations for 1000 error simulations. Pollutant labels are in order of increasing population-weighted semivariance.

To assess a range of error types, simulations were generated with values of $\sigma_{\ln Z}/\sigma_{\ln Z^*}$ ranging from that of error type C to that of type B (eq. 4.10) for the case of an amount of error representative of CO ($\overline{\gamma'} = 0.411$). Epidemiologic model results for RR attenuation are shown in Figure 4.7. On a per unit of measurement (ppm) basis, RR attenuation increased from -24% (i.e. a bias away from the null) for type B error to 85% for type C error. On a per IQR basis, RR attenuation increased from 28% for type B error to 85% for type C error. It is interesting to note that for $\sigma_{\ln Z}/\sigma_{\ln Z^*} = 0.74$ the error ($Z - Z^*$) is independent of Z (i.e. $R(Z - Z^*, Z) = 0$) and the RR per unit attenuation is 0. This is the expected result when error is the Berkson type on an unlogged basis.

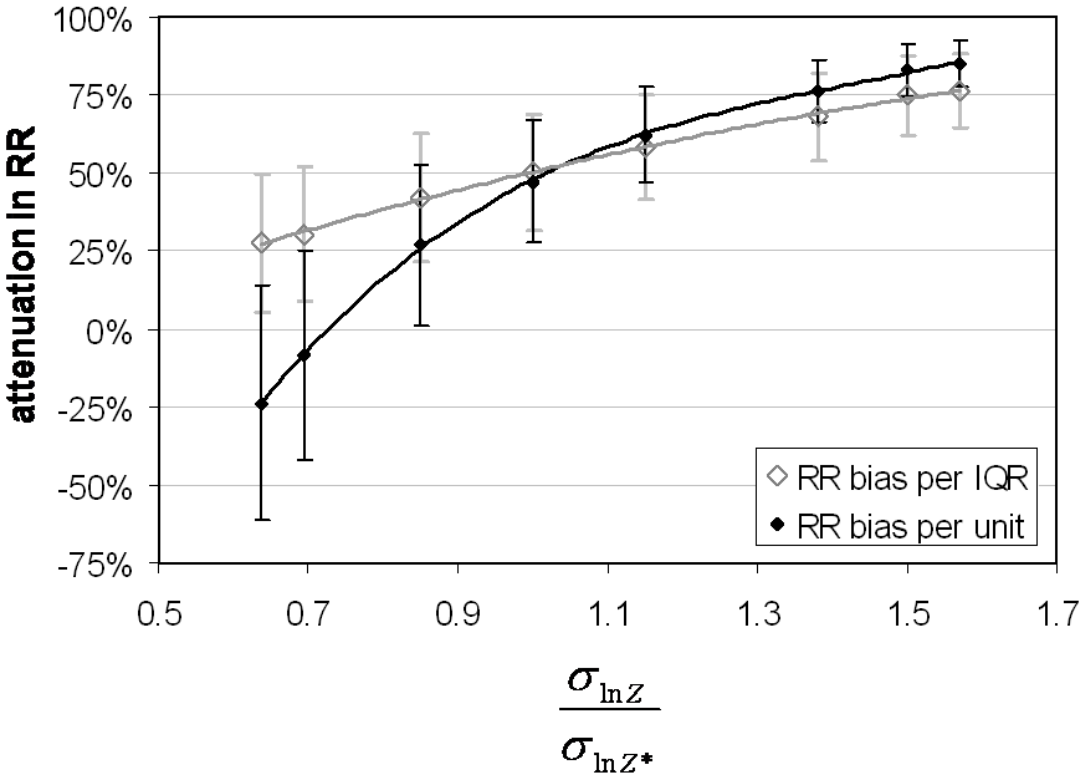


Figure 4.7. Percent attenuation in risk ratio per unit of measurement (ppm) and per IQR for CO error simulations ($\overline{\gamma'} = 0.411$) with incremental changes in error type ranging from type B ($\sigma_{\ln Z} / \sigma_{\ln Z^*} = 0.65$) to type C ($\sigma_{\ln Z} / \sigma_{\ln Z^*} = 1.55$). Bars denote standard deviations for 1000 simulations.

4.4 Discussion

The results demonstrate that error type affects the reduction in significance as well as the RR estimate in the epidemiologic analysis. Moreover, the results demonstrate a profound effect of error type on the RR estimate per unit of measurement. The RR per unit of measurement estimate is increased by the presence of type B error; that is, there is a bias away from the null. To better understand these results, we estimate the attenuation in the effect estimator β (eq. 4.11) in the absence of confounders from the first-order linear regression coefficient (m) of error (Z-Z^*) versus Z as follows.

$$\frac{\beta}{\beta^*} = 1 - m \tag{4.15}$$

For RR estimates near 1 (i.e. β values near 0) as is the case in this study, the predicted attenuation in RR is approximately given as follows.

$$\text{RR per unit attenuation} \approx m \tag{4.16}$$

$$\text{RR per IQR attenuation} \approx 1-(1-m)\frac{IQR}{IQR^{*}} \tag{4.17}$$

Epidemiologic model results are compared with the predictions of eq. 4.16 and eq. 4.17 for all pollutants and both error types (Figure 4.8). The degree to which the epidemiologic results differ from these predictions likely indicates the degree to which confounding variables are affecting results. As shown by the 1:1 line in Figure 4.8, there is strong agreement between the attenuation predicted by analysis of the error model results (i.e. *m* and IQR) and that obtained from the epidemiologic model.

In this study, in which quantification of error is based on the variability between monitors, error due to spatial variation is much greater than error due to instrument imprecision, particularly for primary air pollutants (Goldman et al. 2010). Conceptually, therefore, we speculate that this error is more likely of the Berkson type, with true values varying randomly about a population-weighted average represented by the base case. If spatial error is best described by the Berkson-like type defined on a log basis (our error type B) and the mean of the measurements is the same mean as the true values, we estimate there to be a 24% to 34% attenuation in RR per IQR estimates (Figure 4.6, right panel), and a 19% to 31% bias away from the null in RR estimates on a per unit of measurement basis (Figure 4.6, left panel), for the primary pollutants studied (SO_2, NO_2/NO_x, CO, and EC) when using a population-weighted average as the exposure metric. For the secondary pollutants and pollutants of mixed origin (O_3, SO_4, NO_3, NH_4, $PM_{2.5}$, OC, and PM_{10}), we estimate a 5% to 15% attenuation in RR per IQR estimates and a 2% to 9% bias away from the null in RR estimates on a per unit of measurement basis. We are currently investigating different methods for estimating actual error type based on

simulated pollutant fields trained to have all of the characteristics, including the pattern of spatial autocorrelation, expected of true pollutant fields.

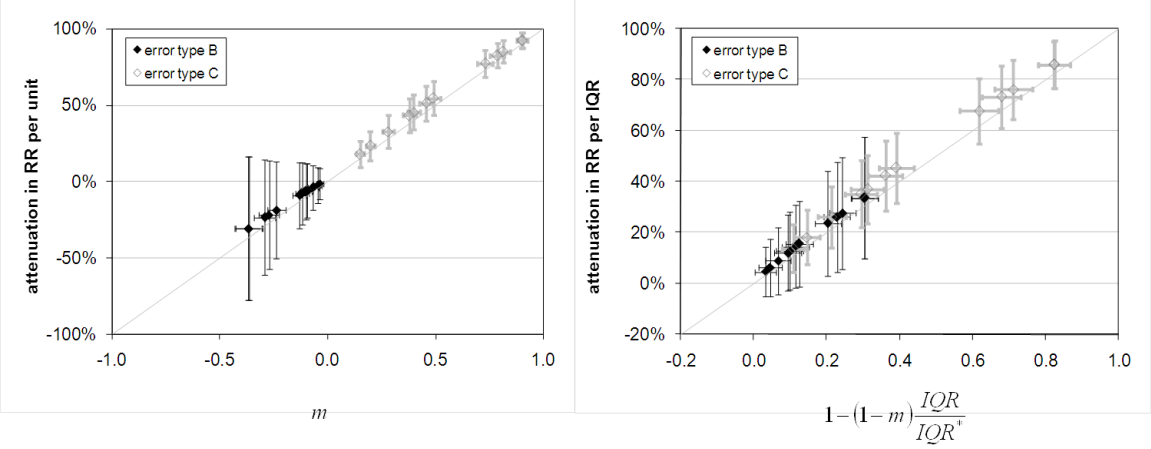


Figure 4.8. Attenuation in the risk ratio per unit of measurement (left panel) and per IQR (right panel) due to the introduction of measurement error, modeled both as type B and type C error. Ranges denote standard deviations for 1000 simulations. One-to-one line is also shown.

This study addresses error between measured and true ambient concentrations. Our results are consistent with previous finding that suggest that Berkson error, as defined on an unlogged scale (additive), produces no bias in the effect estimate (Armstrong 1998; Zeger et al. 2000) as shown in Figure 4.7; however, Berkson-like error defined on a log basis (multiplicative) can lead to risk ratio estimates per unit increase that are biased away from the null (although with a reduction in significance). Thus, the direction and magnitude of the bias are functions of error type. With the multiplicative error structure used here in conjunction with a linear dose response, large "true" values of air pollution would likely be underestimated, resulting in an overestimation of pollution health effects. We have shown how multiple air pollution measurements over space can be used to quantify the amount of error and provide a strategy for evaluating impacts of different types of this error. The results suggest that estimating impacts of measurement error on health risk assessment are particularly important when comparing results across primary and secondary pollutants as the corresponding error will vary widely in both amount and type depending on the degree of spatial variability. These results are

suggestive of error impacts one would have from time-series studies in which a single measure, such as the population-weighted average, is used to characterize an urban or regional population exposure. The methodology used here can be applied to other study areas to quantify this type of measurement error and quantify its impacts on health risk estimates.

4.5 Conclusion

Health risk estimates of exposure to ambient air pollution are impacted by both the amount and the type of measurement error present, and these impacts vary substantially across pollutants. By modeling combined instrument imprecision and spatial variability over a range of error types, we are able to estimate a range of effects of these sources of measurement error, which are likely a mixture of both classical and Berkson error types. This study demonstrates the potential impact of measurement error in an air pollution epidemiology time-series study and how this impact depends on error type and amount. Whereas differences in the amount of error across pollutants have been characterized, work is ongoing to develop methodologies for assessing how error type differs across pollutants.

CHAPTER 5

QUANTIFICATION AND IMPACTS OF SPATIAL VARIABILITY ERROR IN TIME-SERIES STUDIES OF AMBIENT AIR POLLUTION: A SIMULATION STUDY

Abstract

In recent years, geostatistical modeling has shown promising results for gaining insight for health studies looking at ambient air pollution levels. In this study, distributions of daily ambient concentrations are modeled over space and time for 12 air pollutants. Simulated pollutant fields are produced for a 6-year time period over a 20-county metropolitan Atlanta area. Simulations take into account the temporal and spatial autocorrelation structure of ambient pollutants, as well as trends for season, day-of-week and distance from urban center. The simulated "true" ambient concentrations are then utilized to generate simulated monitor data that contain instrument error, including error associated with instrument imprecision as well as with the number and placement of monitors. Error type is characterized, and the impact of error on the epidemiologic analysis is predicted. Measurement error due to spatial variability alone was found to be largely Berkson, suggesting reductions in significance but minimal risk attenuation in time-series risk estimates due to this error source.Total measurement error, consisting of spatial variability error and error associated with instrument imprecision and location, however, results in substantial attenuation of the risk estimate, particularly for primary pollutants.

5.1 Introduction

Measurement error is inherent in time-series epidemiologic studies of air pollution that rely on ambient monitor data. Instrument error and, to a greater degree, exposure misclassification due to the spatial variability have been shown to bias effect estimates in large population studies (Chen et al. 2007; Goldman et al. 2010; Sarnat et al. 2010; Wilson et al. 2007). Time-series studies that rely on central monitor data have been criticized for uncertainty related to exposure measurement errors and the substantial variation present in some air pollutant measures (Dominici et al. 2006). Both error amount and error type affect health risk estimates and significance. Error type (e.g. classical and Berkson) has been demonstrated to modify the extent to which measurement error attenuates health effect estimates (Armstrong 1998; Goldman et al. 2011; Sheppard et al. 2005; Zeger et al. 2000). Classical error is that in which the measurement varies randomly about the true exposure and produces a biased effect estimate. In contrast, Berkson error is that in which the true exposure varies randomly about the measurement, such as might be the case if a population-weighted average exposure is used to characterize ambient exposure. Purely Berkson-type error is expected to decrease significance of an association but will yield an unbiased effect estimate (Armstrong 1998; Zeger et al. 2000). Because the distribution of true concentrations cannot be known with certainty, assessment of error type for a given dataset is challenging.

Increasingly, advanced spatial modeling techniques are being employed in order to gain insight on the distribution of true ambient concentrations (Jerrett et al. 2010). Several studies have developed methods of simulating air pollutants, taking into account both spatial and temporal characteristics of concentrations (Nunes and Soares 2005; Sahu and Mardia 2005); however, few studies have used such simulations to assess the amount and type of measurement errors in time-series studies (Peng and Bell 2010). Gryparis and coauthors (2009) used a smoothing method of spatial measurement error modeling to explore the relative performance of use of different exposure metrics in a study of

particulate matter (PM) and birth weight in greater Boston. They suggest that exposure simulation can inform on the relative uncertainties associated with use of different exposure metrics. Fuentes et al. (2006) utilized a multivariate spatial regression to model the spatial structure of concentrations in order to quantify uncertainties in an association between mortality and fine PM and components. With respect to time-series studies in particular, Peng and Bell (2010) estimated county-wide average concentrations to assess spatial misalignment error and apply statistical methods to obtain adjusted health risk estimates in a time-series study of PM components and hospital admissions for cardiovascular disease (CVD). Lee and Shaddick (2010) investigated the impacts of spatial variation, monitor placement and measurement error by jointly modeling pollutant concentrations and health data using a Bayesian spatio-temporal model and found that in areas where a large number of monitoring sites are available, pollution surface modeling may provide better health effect estimates, particularly for more spatially varying species.

Although spatial and temporal properties of air pollutants can bias health association effect estimates, current study designs typically do not consider such properties in time-series analysis (Peng and Bell 2010). In our previous work (Chapters 3 and 4), the lack of spatial autocorrelation of pollutant concentrations and instrument error were shown to lead to substantial reductions in significance and potential attenuations in risk estimates (Goldman et al. 2010) with error type affecting the amount of attenuation in risk estimates (Goldman et al. 2011). These studies did not, however, specify the type of error actually present in ambient observations or account for the spatial heterogeneity in pollution levels. Moreover, while there has been substantial discussion of the potential impact of error type, few studies have attempted to determine the error type of air pollutant monitoring data in a time-series setting. To do this, the relationship between the true ambient and the measured ambient used in the health study should be understood and quantified. Emissions-based models of ambient air pollution, such as the Community Multi-scale Air Quality (CMAQ) modeling system, observation-based interpolation

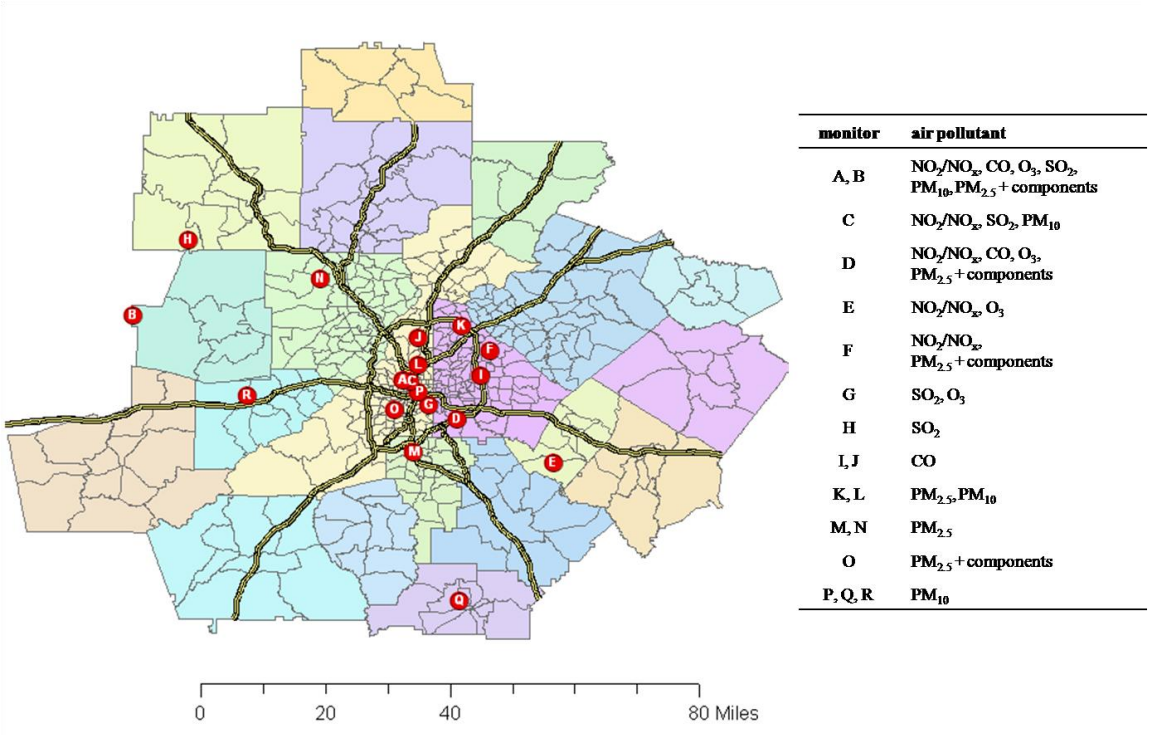
methods, such as kriging (Mulholland et al. 1998), and hybrids of these methods (Kaynak et al. 2009; Mendoza-Dominguez and Russell 2001) are able to capture many characteristics of ambient concentrations at high spatial and temporal resolution; however, these approaches fail to capture the lack of spatial autocorrelation of ambient levels that is particularly relevant in time-series studies. In this work, we create simulated ambient air pollution fields that have the desired spatial and temporal distribution properties found to be present in actual ambient monitor data for 12 air pollutants. For each pollutant time-series, six properties were modeled: temporal autocorrelation, spatial autocorrelation, mean, standard deviation, seasonal trend, and day-of-week trend. Using these "true" ambient air pollution fields, monitor data is simulated and the amount and type of measurement error present for different exposure metrics (i.e. central monitor data and various monitor averages) is assessed. Finally, the impact of measurement error on health risk estimates is predicted based on an analysis of the dependence of error on the simulated measurement. Work is ongoing to assess measurement error impact by simulating health outcome data (i.e. emergency department visits for acute illnesses) based on a specified association with the simulated true ambient and performing time-series analyses using the different exposure metrics.

5.2 Methods

5.2.1 Air Quality Data

To assess spatio-temporal trends in air pollutant concentrations, daily measures of ambient monitor data for the 20-county study area for a 6-year period (1999-2004) were analyzed for 12 ambient air pollutants: 1-hr max NO_2, 1-hr max NO_x, 8-hr max O_3, 1-hr max SO_2, 1-hr max CO, 24-hr PM_{10} mass, 24-hr $PM_{2.5}$ mass, and 24-hr $PM_{2.5}$ components sulfate (SO_4), nitrate (NO_3), ammonium (NH_4), elemental carbon (EC) and organic carbon (OC). Data was obtained from 3 monitoring networks: the US EPA's Air Quality System (AQS), including State and Local Air Monitoring System and Speciation Trends

Network for $PM_{2.5}$ component measurements; the Southeastern Aerosol Research and Characterization Study (SEARCH) network (Hansen et al. 2003), including the Atlanta EPA supersite at Jefferson Street (Solomon et al. 2003b)**;** and the Assessment of Spatial Aerosol Composition in Atlanta (ASACA) network (Butler et al. 2003). While some differences exist between measurement methods used by the monitoring networks, these discrepancies have been discussed in detail elsewhere (Goldman et al. 2010) and are expected to have a negligible impact on this analysis. Monitor site locations are shown in Figure 5.1.



monitor	air pollutant
A, B	NO_2/NO_x, CO, O_3, SO_2, PM_{10}, $PM_{2.5}$ + components
C	NO_2/NO_x, SO_2, PM_{10}
D	NO_2/NO_x, CO, O_3, $PM_{2.5}$ + components
E	NO_2/NO_x, O_3
F	NO_2/NO_x, $PM_{2.5}$ + components
G	SO_2, O_3
H	SO_2
I, J	CO
K, L	$PM_{2.5}$, PM_{10}
M, N	$PM_{2.5}$
O	$PM_{2.5}$ + components
P, Q, R	PM_{10}

Figure 5.1. Map of 20-county metropolitan Atlanta study area. Census tracts, expressways, and ambient air pollutant monitoring sites are shown.

5.2.2 Characterization of Air Pollutant Temporal and Spatial Distributions

To produce simulations of "true" daily ambient concentration fields, ambient monitor data were used to assess the following distributional features: temporal autocorrelation, spatial autocorrelation, mean, standard deviation, seasonal trend, and day-of-week trend. In order to capture the desired trends in addition to the spatial and

temporal autocorrelation structure, an iterative simulation process was performed. The resulting ambient pollutant concentration fields do not simulate actual ambient pollutant concentration fields on any particular day, but rather provide a time-series of fields that possess properties of the actual time-series of fields that are thought to be important for assessing measurement error.

Short-term temporal autocorrelation is present in ambient air pollutant data due to meteorological events occurring on time-scales of a few days to a week. Correlations of data from each monitor were calculated for one to seven day lags (Figure C.1). In general, the temporal autocorrelation trend with increasing lag was similar for urban, suburban and rural monitors, so this short-term temporal autocorrelation was characterized as being independent of location. It was also similar across pollutants; 1-day and 2-day lag Pearson correlation coefficients of 0.59 ± 0.13 and 0.33 ± 0.17, respectively. Secondary pollutants tended to have greater levels of short-term temporal autocorrelation and these correlations persisted over longer lag periods. Longer term temporal autocorrelation was also observed (Figure C.2), but this was simulated via seasonal trend modeling.

Spatial autocorrelation is present in ambient air pollutant data due to the distribution of emission sources and transport phenomena. The spatial autocorrelation of pollutants was assessed using correlograms (Figure C.3) constructed using all available data. Primary pollutants, i.e. those largely emitted directly to the atmosphere such as CO, SO_2 and EC, were observed to have much less spatial autocorrelation than secondary pollutants, i.e. those largely formed in the atmosphere such as O_3, NO_3 and SO_4. Pollutants of mixed origin, e.g. $PM_{2.5}$ total mass and OC, were found to have intermediate levels of spatial autocorrelation. Population-weighted average spatial autocorrelations were calculated for each pollutant (Table C.1). Values ranged from 0.901 for O_3 to 0.208 for SO_2.

The distributions of most air pollutant concentrations were found to be best modeled as lognormal. Primary pollutants tend to be spatially heterogeneous, with means and standard deviations decreasing with increasing distance from the urban center. Secondary pollutants tend to be spatially homogeneous with little change in their distribution over space. The mean and standard deviation of air pollutant log concentrations were modeled as a linear function of distance from the urban center up to 60 km using data from all available monitors; beyond 60 km from the urban center the mean and standard deviation were fixed at the rural background levels. Measured means and standard deviations are shown as functions of distance from the urban center in Figures C.4 and C.5, respectively.

Day-of-week and seasonal trends in air pollutant concentrations were also characterized. These trends result from variation in emissions by day-of-week and season and by variation in meteorological parameters affecting mixing rates and formation and removal rates by season. To assess the day-of-week patterns, daily pollutant concentrations from each monitor were normalized by the mean, grouped categorically by day-of-week, and averaged. The day-of-week trends were found to be similar at different monitor locations. For seasonal trends, pollutant concentrations from each monitor were normalized by the mean, averaged by day-of-year, and modeled using fourth order polynomial functions under the constraints that the value and slope on day 366 are the same as those on day 1. Seasonal trends were found to differ between monitors in urban, suburban, and rural locations; therefore, the polynomial seasonal trend was modeled to change as a linear function of distance from the urban center up to 60 km. Beyond 60 km, the seasonal trend was fixed at the rural background trend. Results of the characterization of day-of-week and seasonal trends are shown in Figures C.6 and C.7, respectively.

5.2.3 Simulation of Ambient Pollutant Fields

Daily air pollutant fields were generated for the 20-county Atlanta region (16,000 km^2) at a 5-km resolution for a 6-year period (2,192 days) for the 12 ambient air pollutants based on the characteristics described above. Simulations were produced via a two-step process. First, the direct sequential simulation method (Soares 2001) in the Stanford Geostatistical Modeling Software (SGeMS) (Remy 2005) was used to generate normalized fields (eq. 5.1) with the desired short-term temporal and spatial autocorrelation (Figures C.1 and C.3, respectively).

$$\chi_{ij}^{*} = \frac{\ln C_{ij} - \mu_{\ln C_j}}{\sigma_{\ln C_j}} \sim \mathrm{N}(0,1) \tag{5.1}$$

Here, χ_{ij}^{*} is the normalized "true" pollutant level on day *i* at location *j*, C_{ij} is the concentration on day *i* at location *j*, $\mu_{\ln C_j}$ is the log concentration mean over all days at location *j*, and $\sigma_{\ln C_j}$ is the log concentration standard deviation over all days at location *j*. Thus, at each location *j*, χ_{ij}^{*} has a mean of zero and a standard deviation of one. SGeMS was used in this application to provide the same autocorrelation in two spatial dimensions and a different autocorrelation in a third (temporal) dimension.

The second step in generating pollutant field simulations was denormalization to yield concentration fields with the desired means, standard deviations, day-of-week trends, and seasonal trends. This was achieved by inverting eq. 5.1 and applying factors to achieve the desired day-of-week and seasonal trends (eq. 5.2).

$$Z_{ij}^{*} = \alpha_{wk}\alpha_{yr} \exp\left(\chi_{ij}^{*}\sigma_r + \mu_r\right) \tag{5.2}$$

Here, Z_{ij}^{*} is the "true" concentration on day *i* at location *j*, μ_r is the log concentration mean modeled as a function of distance from urban center *r*, σ_r is the log concentration standard deviation modeled as a function of *r*, α_{wk} is the day-of-week factor, and α_{yr} is the season factor modeled as a function of *r*. An iterative process was performed in order

to preserve the desired levels of appropriate spatial and temporal autocorrelation structure after the concentrations were denormalized and distribution trends were added. Values for each of the model parameters, μ_r, σ_r, α_{wk} and α_{yr}, for each pollutant are provided in the Appendix (Tables C.2-C.5).

Because many time-series study designs use a single daily metric to represent ambient pollutant exposure for a population, a population-weighted average time-series of the true ambient concentration fields was computed to quantify the amount of measurement error due to spatial variability alone, assuming that it is perfectly measured. Population at the census tract level from the 2000 US Census was utilized, weighting the concentration value for each grid cell by the fraction of population residing within that cell. Population was assumed uniformly distributed within census tracts and area-weighted fractions of population were assigned to respective grid cells for census tracts that fell within more than one cell.

5.2.4 Simulation of Monitor Data and Calculation of Exposure Metrics

Simulated time-series were generated at ambient monitor locations for each pollutant to represent measured data by adding classical-like error to the true ambient level in amounts consistent with information obtained from collocated instruments. The method for characterizing instrument error developed in previous work (Goldman et al. 2010; Goldman et al. 2011) was used here. Error was then added on a log scale, yielding simulated monitor data with the same log mean and standard deviation as the true ambient time-series at that location.

The simulated monitor data were used to compute the following exposure metrics: central monitor, unweighted average of monitor data, population-weighted average of monitor data, and area-weighted average of monitor data. The first exposure metric examined was the simulated Jefferson Street time-series, as many time-series studies utilize measurements from a central monitor to assess exposure (Wilson et al. 2005). A

time-series of the daily averages of simulated data for all monitors for each pollutant was computed as a second exposure metric. The number of monitors simulated ranged from five for CO to twelve for $PM_{2.5}$ mass. A population-weighted average time-series was computed from the simulated monitor data using census tract population from the 2000 Census and a previously developed spatial interpolation method (Ivy et al. 2008) to compute concentrations between simulated monitor values. Lastly, an area-weighted average was computed using the simulated monitor data. These four time-series represent different exposure metrics that have been used in time-series studies and their relative representativeness is assessed.

Measurement error, ε, was calculated for each exposure metric as the difference between the simulated measurement for that exposure metric (Z) and the simulated true ambient concentration field (Z_{ij}^{*}). Population-weighted Pearson correlation coefficients were computed between ε and the "measured" time-series (Z) and between ε and the "true" concentrations (Z_{ij}^{*}). For Berkson error, the expected value of $R(\varepsilon,Z)$ is zero, whereas for classical error the expected value of $R(\varepsilon,Z^{*})$ is zero. Finally, an expected amount of bias in risk ratio estimates due to measurement error for each pollutant was calculated as the slope of ε versus Z based on our previous findings (Goldman et al. 2011).

5.3 Results

5.3.1 Ambient Concentration Field Simulation

Ambient concentration fields were simulated to have six features observed in monitor data and thought to be important in assessing measurement error impacts in time-series health studies: temporal autocorrelation, spatial autocorrelation, mean, standard deviation, seasonal trend, and day-of-week trend. Using SGeMS, simulations were generated to have the level of temporal and spatial autocorrelation observed in the

monitor data. The spatial autocorrelation of ambient concentrations, as characterized by correlograms of monitoring site data, is well captured by the simulated time-series at both urban (Figure 5.2) and rural (Figure C.8) locations, suggesting that the simulations have an appropriate correlation structure at all locations in space. Observed short-term temporal autocorrelation is approximated in the simulations for lags up to 14 days (Figure 5.3). Long-term temporal autocorrelation (over an entire year) was also captured by the modeling of seasonal trends observed in the monitoring data at both urban and rural locations (Figure C.9). The mean and standard deviation as functions of distance from the urban center were also captured well in the simulations (Figures C.10 and C.11, respectively). Day-of-week trends were modeled for pollutants where significant trends were observed in the monitor data (Figure C.12). Seasonal trends observed in the monitor data are well captured by the simulations at both urban and rural locations (Figure C.13).

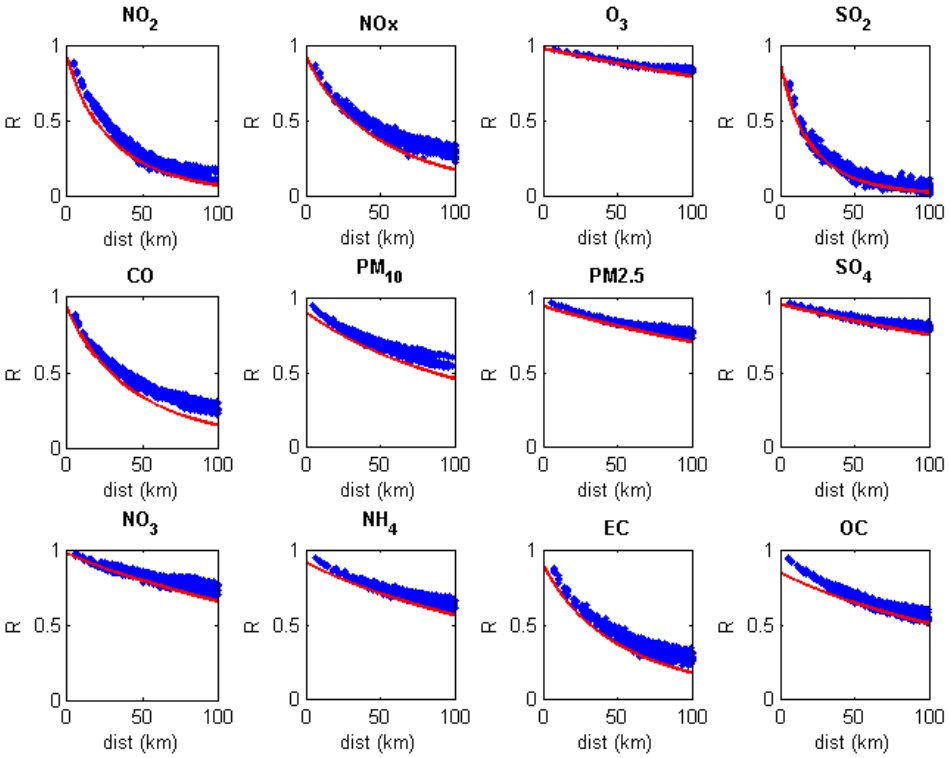


Figure 5.2. Spatial correlogram model regressed from monitoring site data (red) and correlation of simulated time-series at all grid cells with the time-series at the center grid cell (blue).

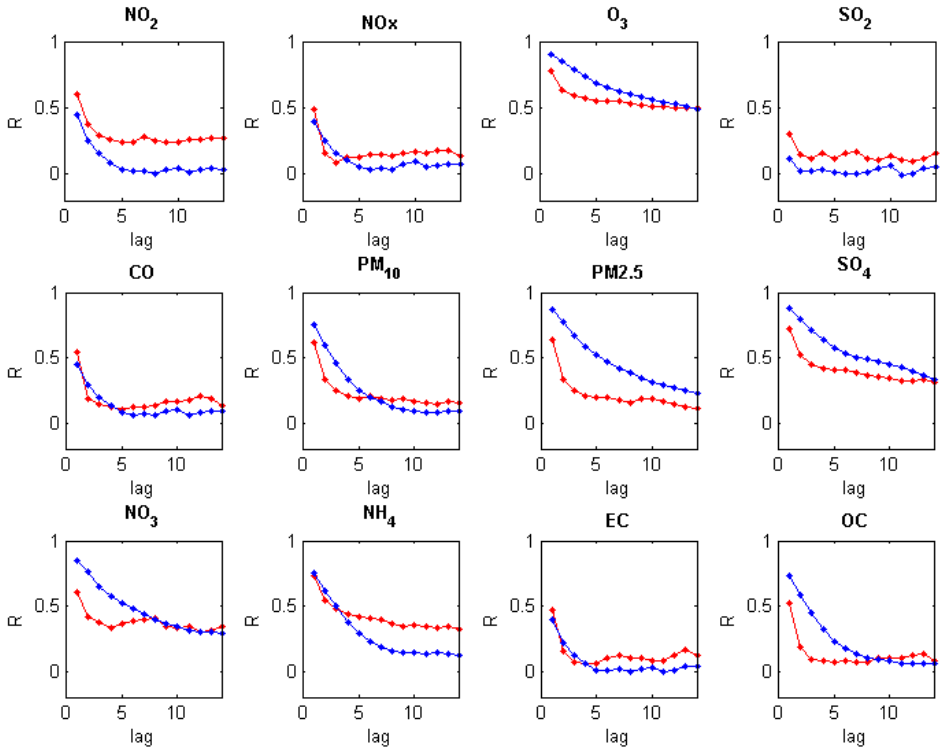


Figure 5.3. Short term temporal autocorrelation of measurements at the central monitoring site (red) and of the simulated time-series at the central grid cell (blue).

5.3.2 Monitor Data and Exposure Metric Simulation

Time-series of monitor data were simulated at the grid cells where actual monitors are located. The amount of error was determined from the correlation between data from collocated instruments for each pollutant, and the type of error added was confirmed to be classical-like (i.e. classical on a log basis); results are provided in Table C.6.

From the simulated monitor data, four alternative exposure variables were computed: central monitor, unweighted monitor average, population-weighted monitor average and area-weighted monitor average. Means and standard deviations for these metrics are provided in Table 5.1, with those of the true population-weighted average also listed for comparison. The central monitor metric yields the largest mean and standard deviation, particularly for primary pollutants which have much higher levels in the urban center than in the rural surroundings. The population-weighted metrics based

on monitor data have a mean and standard deviation most similar to the true population-weighted average values.

Table 5.1. Mean and standard deviation of different exposure metrics and of the true population-weighted average.

	central monitor		unweighted average		population-weighted avg		area-weighted avg.		true population-weighted avg	
pollutant	mean	st dev	mean	st dev	mean	st dev	mean	st dev	mean	st dev
NO_2 (ppb)	44.5	20.3	29.4	12.6	23.5	11.5	15.4	8.8	23.9	10.2
NO_x (ppm)	0.117	0.118	0.065	0.057	0.046	0.041	0.025	0.021	0.048	0.039
O_3 (ppb)	44.9	27.9	44.8	25.3	43.9	25.3	44.1	25.0	44.6	24.5
SO_2 (ppb)	15.4	22.8	12.8	12.3	10.3	10.6	8.9	7.9	12.0	8.3
CO (ppm)	1.62	1.29	1.11	0.70	0.79	0.45	0.49	0.22	0.75	0.36
PM_{10} ($\mu g/m^3$)	25.1	14.9	24.0	13.1	22.6	13.1	21.2	12.5	22.8	12.3
$PM_{2.5}$ ($\mu g/m^3$)	18.0	10.0	17.0	8.9	16.1	8.9	15.1	8.9	16.0	8.5
SO_4 ($\mu g/m^3$)	5.35	5.29	5.03	4.67	4.93	4.71	4.67	4.33	4.94	4.56
NO_3 ($\mu g/m^3$)	1.34	1.56	1.15	1.23	1.11	1.19	1.01	1.07	1.12	1.16
NH_4 ($\mu g/m^3$)	2.32	1.85	2.34	1.72	2.24	1.72	2.22	1.64	2.31	1.71
EC ($\mu g/m^3$)	0.90	0.88	0.70	0.54	0.62	0.49	0.50	0.38	0.66	0.48
OC ($\mu g/m^3$)	5.93	4.29	5.33	3.31	5.10	3.28	4.69	2.90	5.13	3.14

5.3.3 Exposure Metric Evaluation: Error Type, Amount, and Predicted Impact on Health Risk Estimation

Having produced simulated true air pollution fields, simulated monitor data and exposure metrics based on the monitor data, we now address the following questions. What type of error is present for each of the metrics? Which metric is most representative of true ambient exposure? How much bias in the health risk estimate is predicted due to measurement error?

We assess error type by calculating population-weighted Pearson correlation coefficients $R(\varepsilon,Z^*)$ and $R(\varepsilon,Z)$ where error, ε, is defined as $Z - Z^*$, Z^* is the true ambient pollutant field, and Z is the exposure metric of interest. A zero value of $R(\varepsilon,Z^*)$ indicates classical error, and a zero value of $R(\varepsilon,Z)$ indicates Berkson error. These correlation results are shown in Table 5.2 for the three exposure metrics, as well as for the true population-weighted average.

Table 5.2. Population-weighted correlations between measurement error and true values, $R(\varepsilon,Z^*)$, and between measurement error and measured values, $R(\varepsilon,Z)$, where the true values, Z^*, are the true ambient concentration fields and the measured values, Z, are the monitor-based metrics (central monitor, unweighted average and population-weighted average) and the true ambient population-weighted average.

	central monitor		unweighted average		population-weighted avg		area- weighted avg		true population-weighted avg	
pollutant	$R(\varepsilon,Z^*)$	$R(\varepsilon,Z)$	$R(\varepsilon,Z^*)$	$R(\varepsilon,Z)$	$R(\varepsilon,Z^*)$	$R(\varepsilon,Z)$	$R(\varepsilon,Z^*)$	$R(\varepsilon,Z)$	$R(\varepsilon,Z^*)$	$R(\varepsilon,Z)$
NO_2	-0.45	0.62	-0.72	0.21	-0.77	0.14	-0.87	-0.04	0.82	0.00082
NO_x	-0.10	0.83	-0.59	0.35	-0.79	0.07	-0.95	-0.28	0.81	0.00078
O_3	0.03	0.40	-0.19	0.13	-0.20	0.12	-0.23	0.09	0.29	0.00058
SO_2	-0.39	0.75	-0.72	0.38	-0.78	0.28	-0.88	0.11	0.86	0.00024
CO	-0.10	0.86	-0.47	0.53	-0.76	0.19	-0.95	-0.20	0.85	-0.00104
PM_{10}	-0.33	0.40	-0.44	0.15	-0.45	0.14	-0.51	0.07	0.52	0.0152
$PM_{2.5}$	-0.11	0.41	-0.28	0.13	-0.28	0.14	-0.29	0.13	0.38	-0.00057
SO_4	-0.06	0.41	-0.27	0.10	-0.26	0.12	-0.45	-0.10	0.34	0.00037
NO_3	0.10	0.59	-0.29	0.15	-0.36	0.08	-0.55	-0.14	0.40	-0.00119
NH_4	-0.39	0.27	-0.46	0.04	-0.46	0.05	-0.53	-0.03	0.46	0.00026
EC	-0.27	0.66	-0.63	0.16	-0.69	0.08	-0.84	-0.15	0.71	0.00003
OC	-0.19	0.55	-0.45	0.13	-0.46	0.12	-0.61	-0.07	0.51	-0.00005

The results show that the error type is neither classical nor Berkson. If a true population-weighted average were available, that is, if the ambient could be perfectly measured at all locations and then averaged via population-weighting, then the type of error due to using a single metric for the entire population would be Berkson, as suggested by the near-zero values of $R(\varepsilon,Z)$ in the rightmost column of Table 5.3. This type of error would result in no bias in the health effect estimate; however, because the ambient is measured imperfectly and in a limited number of locations, the error type when using the alternative metrics based on monitor data is not Berkson and therefore, would result in bias.

One measure of the relative amount of measurement error present is to compare the representativeness of the alternative exposure metrics. Results of the population-weighted correlation of the true ambient fields with the exposure metrics and with the true population-weighted average time-series are listed in Table 5.3. As expected, primary pollutants such as CO, NO_2, NO_x, SO_2 and EC have lower correlations, indicative of greater measurement error than secondary pollutants such as O_3, SO_4, and NO_3. Also as expected, the central monitor metric was less representative of the true ambient than the monitor average metrics. Interestingly, there was little difference between using an unweighted average of monitor values versus a population-weighted average. Finally, it is noted that for primary pollutants most of the error is due to spatial variability as compared to instrument error when using a monitor-based metric. That is, the monitor average data are almost as correlated with the true ambient pollution field as is the true population-weighted average, suggesting that overall there would be little benefit of having more monitors if population whereabouts in the study area are unknown.

Table 5.3. Population-weighted correlations between the true ambient concentration field and different exposure metrics, $R(Z,Z^*)$, where the exposure metrics are the monitor-based metrics (central monitor, unweighted average, population-weighted average and area-weighted average) and the true ambient population-weighted average.

pollutant	central monitor	unweighted average	population-weighted avg	area-weighted avg	true population-weighted avg
NO_2	0.46	0.60	0.59	0.60	0.66
NO_x	0.56	0.66	0.66	0.68	0.71
O_3	0.93	0.95	0.95	0.95	0.96
SO_2	0.29	0.36	0.37	0.37	0.50
CO	0.53	0.63	0.63	0.64	0.69
PM_{10}	0.73	0.82	0.81	0.82	0.85
$PM_{2.5}$	0.87	0.92	0.92	0.92	0.93
SO_4	0.89	0.93	0.93	0.93	0.94
NO_3	0.87	0.91	0.91	0.91	0.92
NH_4	0.79	0.87	0.87	0.87	0.89
EC	0.54	0.67	0.67	0.67	0.72
OC	0.72	0.83	0.83	0.83	0.86

In previous work we showed that the regressed slope, *m*, of measurement error (Z-Z^*) versus measurement (eq. 5.3) is a good predictor of bias in the risk ratio per measurement error, such that *m* is approximately equal to the attenuation in risk ratio per unit (Goldman et al. 2011).

$$\varepsilon_{ij} = mZ_i + b \tag{5.3}$$

Therefore, we calculated a population-weighted value of *m* for each exposure metric and each pollutant, relative to the true population-weighted average, taken here to be Z^*. Results are shown in Figure 5.4. Risk ratio attenuation is predicted to be highest for primary pollutants (CO, SO_2, NO_2, NO_x, and EC) and lowest for secondary pollutants (O_3, NO_3, SO_4, and NH_4). Predicted bias-to-null is greatest when the central monitor data are used, and smallest when the population-weighted average is used.

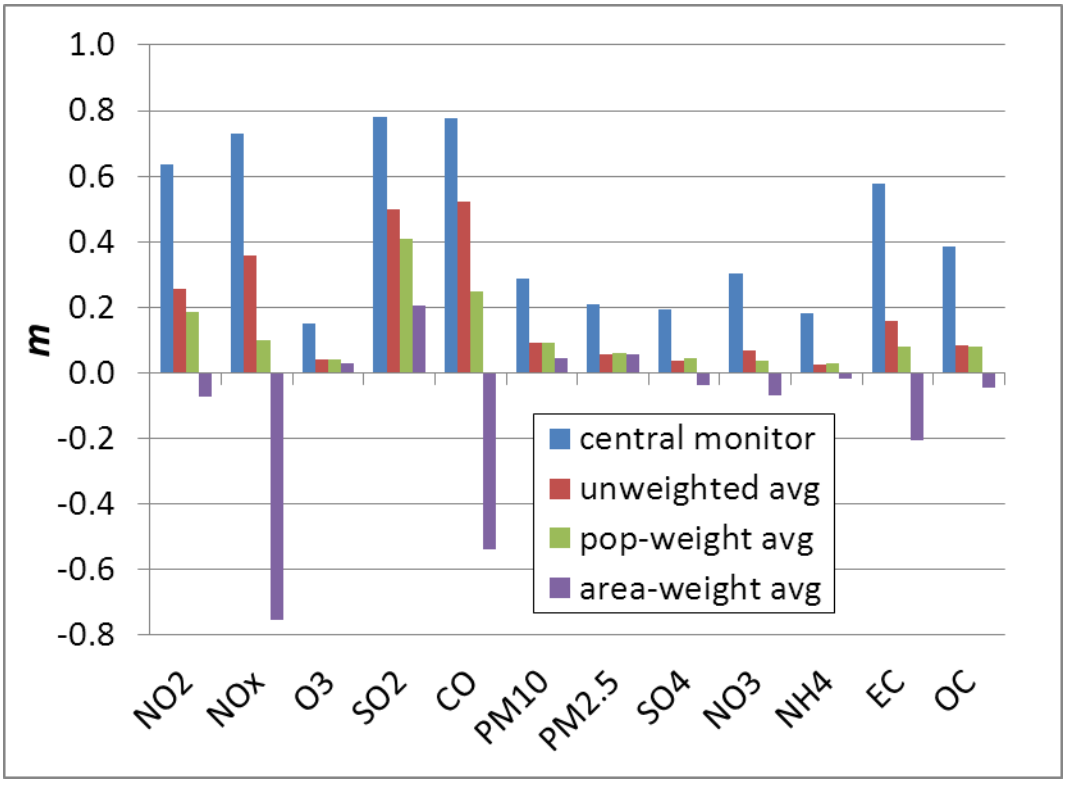


Figure 5.4. Population-weighted slope, *m*, of error versus measurement, where the measurement refers to each exposure metric.

There is no bias due to spatial variability alone because this error is Berkson, as shown by the lack of correlation between error and the true population-weighted average (Table 5.3). Therefore, the bias is introduced when error from the use of instruments is added, where this error is due to measurement imprecision and the number and placement of monitors. Use of an area-weighted metric yields a negative slope, suggesting a bias away from the null hypothesis. The amount of bias is largely due to the difference between the exposure metric temporal variance and the true population-weighted average temporal variance. In Figure 5.5, the predicted bias, *m*, is shown to be a function of the ratio of the standard deviation of the exposure metric used and the standard deviation of the true population-weighted average.

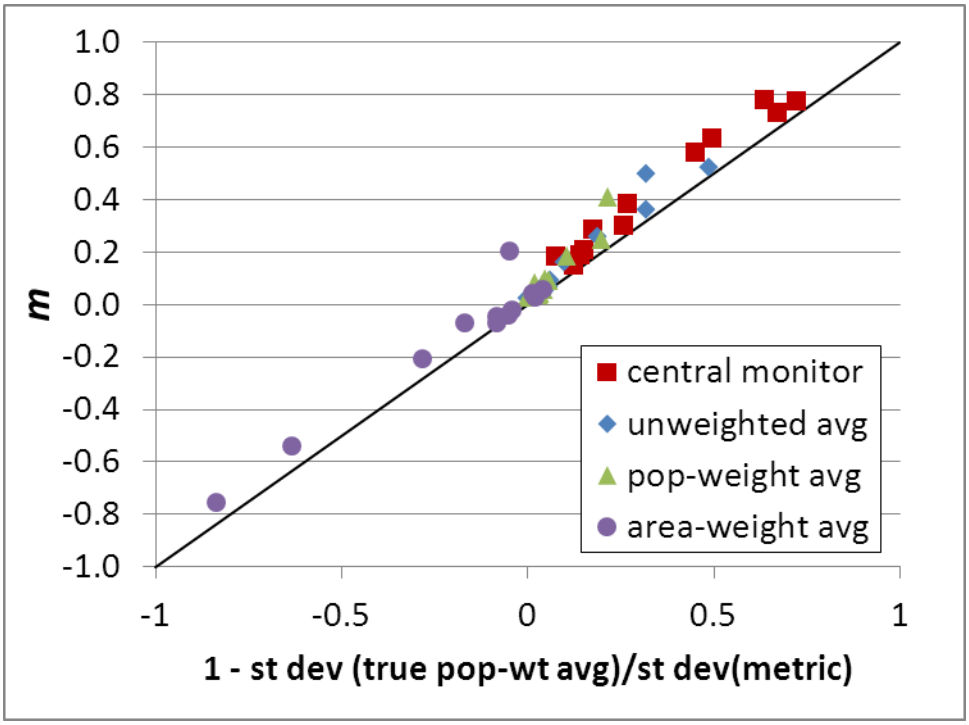


Figure 5.5. Predicted bias in health estimate, *m*, versus one minus the ratio of the standard deviation of the exposure metric used and the standard deviation of the true population-weighted average.

5.4 Discussion

As evident from the lack of correlation between the true ambient concentration fields and the true population-weighted average concentration (Table 5.3), the amount of exposure measurement error in this time-series study of acute health effects and ambient air pollution is largely the result of spatial variability, with much less error derived from the fact that monitor-based exposure metrics are used rather than a true population-weighted average. While this error will result in loss of significance in the health risk assessment, error due to spatial variability alone will not bias the health risk estimate due to it being of the Berkson type (Table 5.2). A true population-weighted average cannot be measured, however, so exposure metrics that are used in health studies also contain instrument error associated with the precision of measurement as well as the number and placement of monitors. While the amount of error due to instrument imprecision is smaller than the amount of error due to spatial variability, particularly for primary

pollutants, this additional error can result in substantial attenuation of the risk ratio estimates. Therefore, the impact of measurement error needs to be assessed in total. In this study of the Atlanta metropolitan area, health risks associated with primary air pollutants are predicted to be attenuated by up to 80% if central monitor data are used, and up to 50% if an average across monitors is used. For secondary pollutants, attenuation can be as low as 15% if the central monitor is used, and as low as 3% if an average across monitors is used.

Use of a population-weighted average of monitor data in the health study is predicted to result in less attenuation of the risk ratio estimate than use of an unweighted average. This result depends on the number and placement of monitors, however. If data from a number of monitors in low-population portions of the study area are used, then a population-weighted average is likely to reduce bias by a larger amount as compared to the bias when using an unweighted monitor average. Population weighting provides a better estimate of the true average variance in exposure, which is the main reason why use of a population-weighted average results in less bias.

Preliminary results for simulation of ED counts without potential confounders and assessment of the associations with the various pollutant metrics were calculated. Figure 5.6 illustrates the percent bias in the epidemiologic effect estimator, β, obtained from regression on the natural log of ED counts versus each pollutant exposure metric. These results agree very well with the predictions shown in Figure 5.4 based on the relationship of error to pollutant exposure metrics. Similar to the relationship between slope and the ratio of the standard deviation of the exposure metric used and the standard deviation of the true population-weighted average show in Figure 5.5, the effect estimator, β is also a function of this ratio in standard deviations (Figure 5.7).

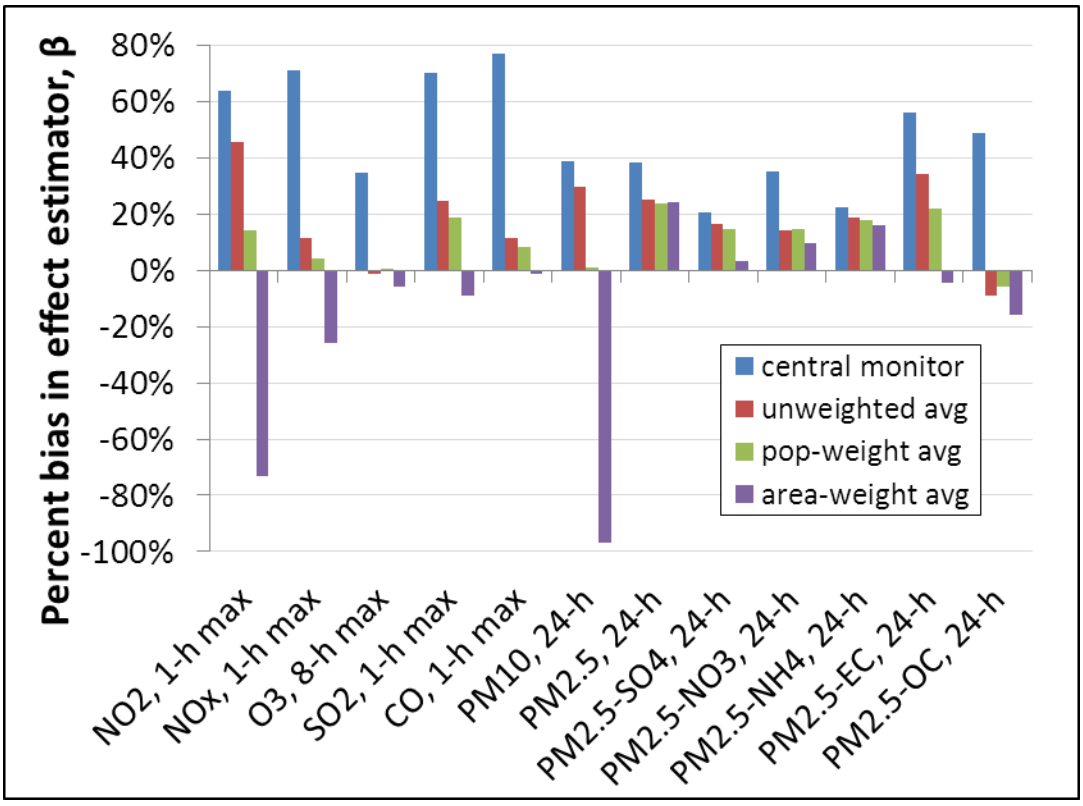


Figure 5.6. Percent bias in effect estimator, β, from regression on simulated emergency department visits versus pollutant exposure metrics.

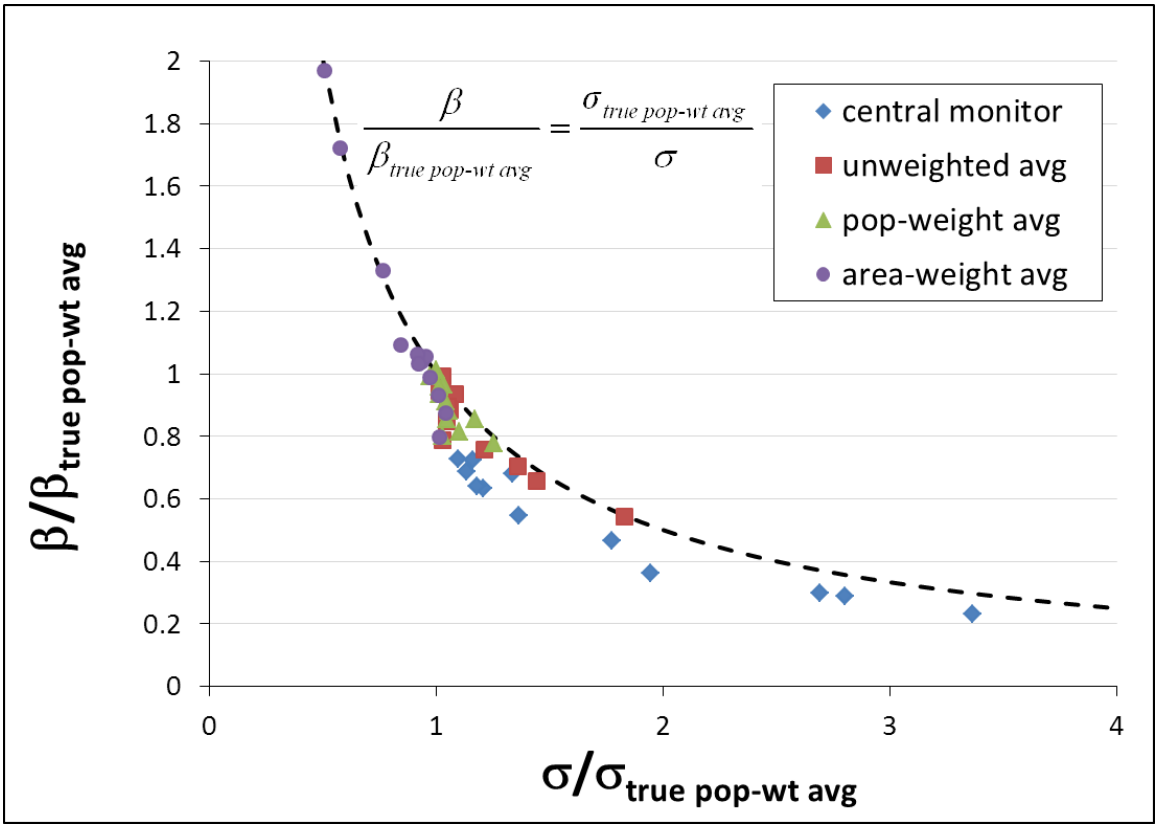


Figure 5.7. The bias in effect estimator, β, relative to that obtained using the true population-weighted average versus the standard deviation in the metric used relative to the standard deviation of the true population-weighted average.

Preliminary results for significance of associations found with use of different metrics were also calculated. Figure 5.8 shows the ratio of the χ^2 values for use of each exposure metric in the epidemiologic model to the χ^2 found using the true population-weighted average metric. The largest reduction in significance, that is, the lowest ratio, was found for use of the central monitor. Use of monitor averages resulted in similar reductions in significance level. The reduction in χ^2 due to use of monitor-based metrics was found to be largely a function of a population-weighted R^2 value between the exposure metric and the simulated true ambient concentration field, suggesting that significance level is strongly influenced by the correlation between the exposure metric used and the true exposure. Future health work will include effects of potential confounders.

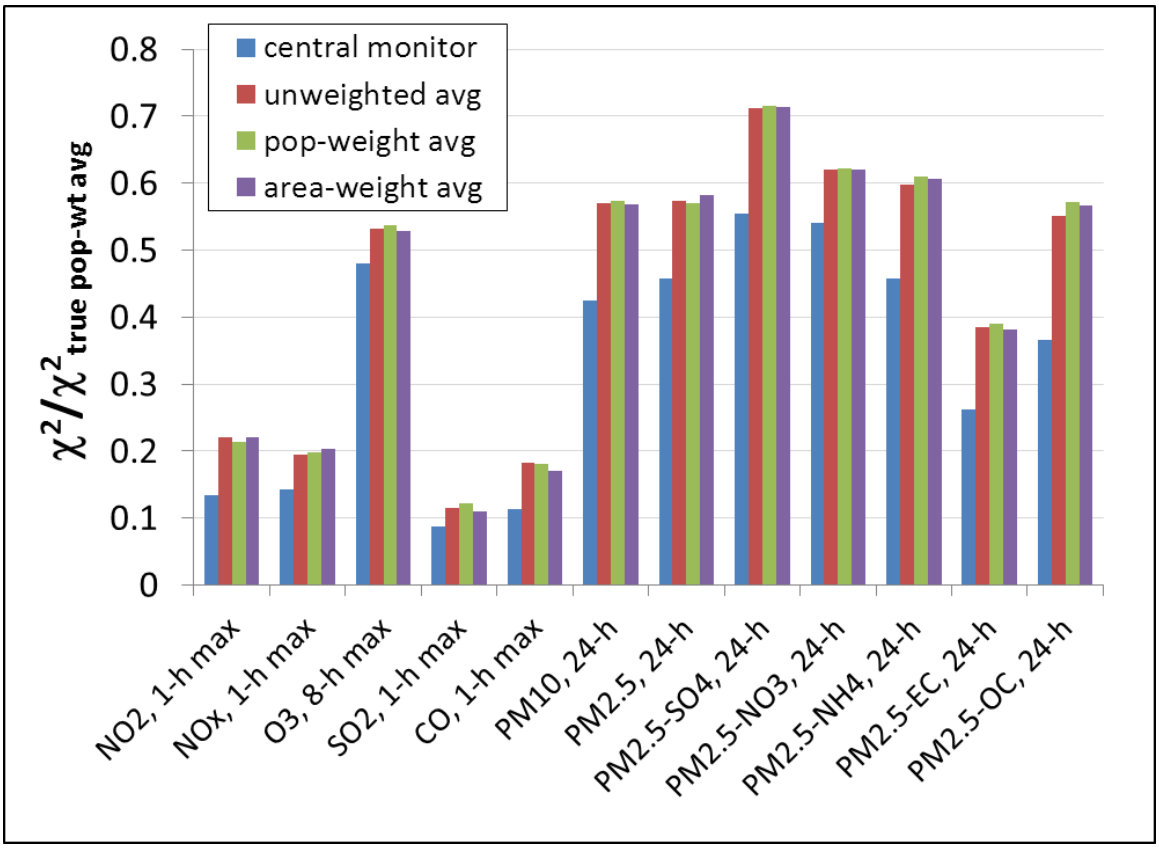


Figure 5.8. Reduction in χ^2 relative to the χ^2 found using the true population-weighted average for each exposure metric.

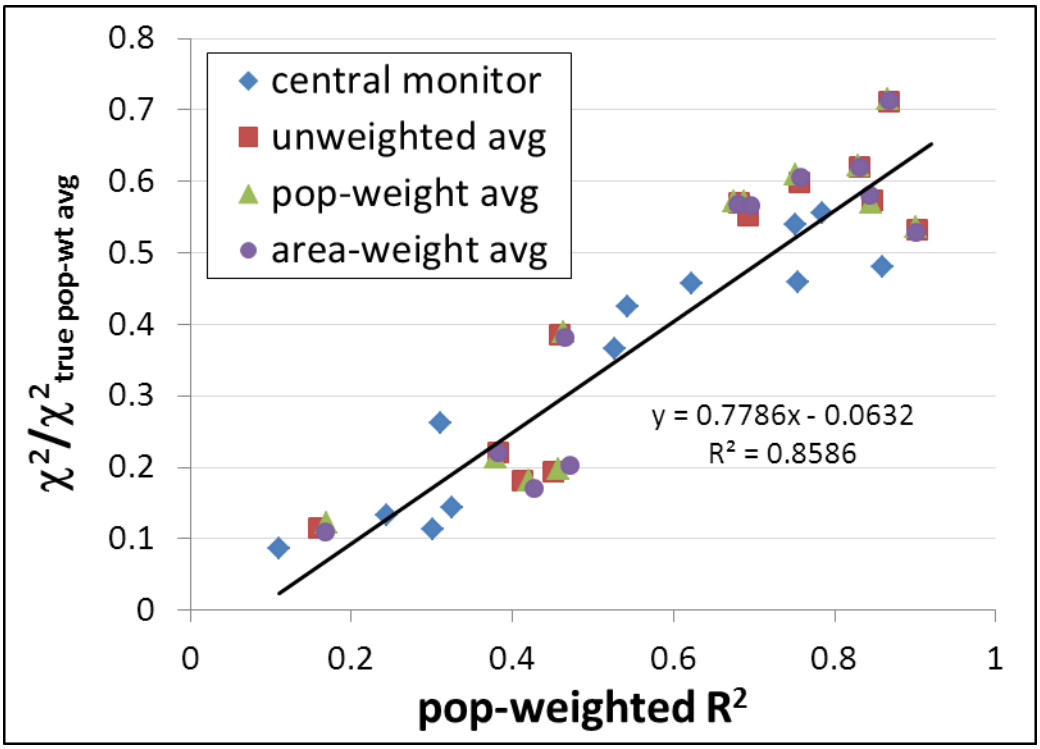


Figure 5.9. Reduction in χ^2 relative to the χ^2 found using the true population-weighted average versus a population-weighted average R^2 value between the simulated true ambient concentration field and each exposure time-series.

The ambient concentration variability modeled here is representative of 'regulatory ambient' variability, that is, the variability expected of outdoor monitors sited to capture ambient pollutant levels used for regulatory purposes. Microscale variability in space and time, such as that which occurs near roadways or near point sources, was not modeled; however, the method presented here could easily be applied for such analyses. There is potential to utilize this procedure to simulate variability in personal exposures, for example, and assess measurement error associated with using ambient monitors rather than personal monitoring. Simulated pollutant fields can also be used to estimate personal exposure in a time-activity model setting, since the concentration fields can be resolved in both time and space at scales commensurate with the scales that air pollution variability has been characterized.

While there is considerable potential for using this methodology for future work, one limitation of the current use of this model should be recognized. A stationary isotropic semivariance model was assumed here. While this simplification is reasonable for many applications, true variance of pollutant concentrations over space and time is likely to have a more complex spatial and temporal variance structure.

The total bias in risk ratio point estimates and reduction in significance level due to total air pollution measurement error as assessed in this work can be used to estimate the health associations that would have been observed if this air pollution measurement error could be effectively removed. Taking the percent bias estimated in this work to be independent of the base case health association, which our simulations have suggested is true, the bias in the observed health associations from a previously published epidemiologic study of cardiovascular disease emergency department visits assessed against Atlanta central monitor measurements (Metzger et al. 2004), can be adjusted to association levels that might have been observed if the measurement error assessed here was removed (Figure 5.10).

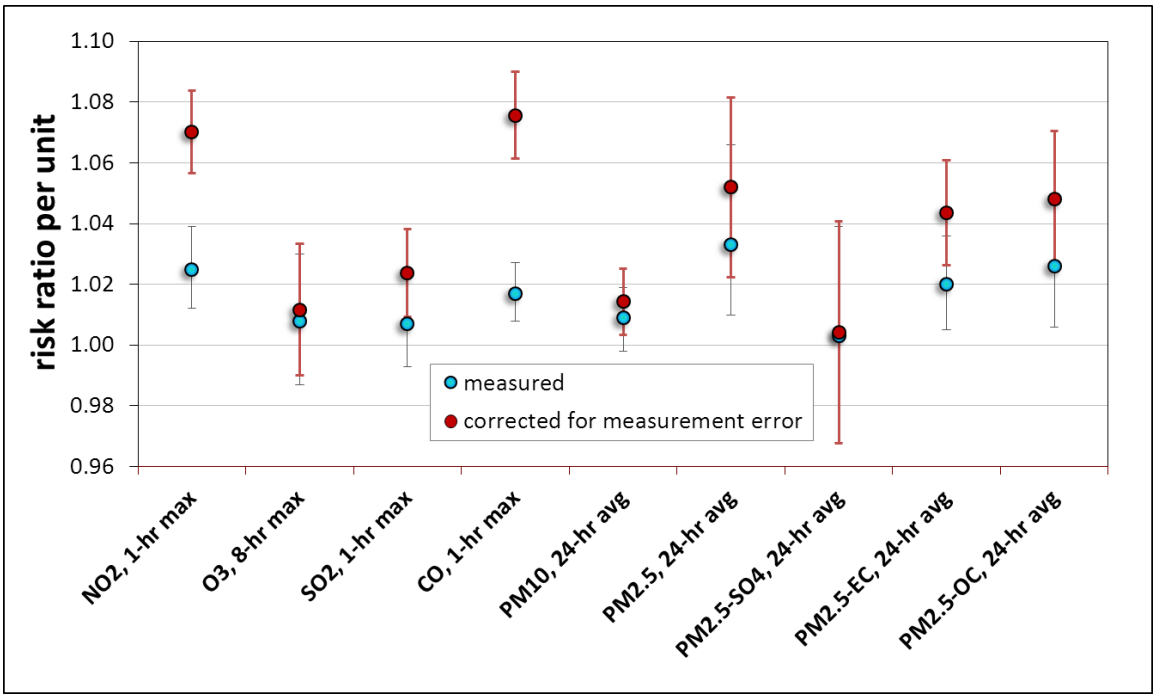


Figure 5.10. Published risk ratio estimates and 95% confidence intervals for cardiovascular disease emergency department visits assessed against Atlanta central monitor measurements adjusted for measurement error.

This work demonstrates a method for simulating ambient air pollutant concentrations over space and time which allows for assessment of the amount and type of error present in a time-series health study. Attenuation in risk ratio estimates was predicted for use of different monitor-based exposure metrics. In ongoing work, the simulations presented here are being coupled with health outcome simulations for use in an epidemiologic model to assess the impact of measurement error on risk estimates and significance levels. Thus, this future work will test the predictions of risk ratio bias presented here and supplement these findings with estimates of reductions in significance.

5.5 Conclusion

Geostatistical modeling of ambient air pollutant concentrations over space and time can provide valuable insights on the amount and type of measurement error present

in time-series epidemiologic studies that use monitor-based exposure metrics. The amount and type of measurement error due to spatial variability of ambient concentrations are assessed and the amount and type of additional error introduced through use of monitor-based ambient exposure metrics is estimated. The amount of measurement error was found to be predominantly that associated with spatial variability. Total measurement error, including that associated with instrument imprecision and the number and placement of monitors, was found to be neither Berkson nor classical in type. As a consequence, this measurement error is predicted to attenuate risk estimates as well as decrease significance of association.

CHAPTER 6

SPATIO-TEMPORAL VARIABILITY OF AMBIENT AIR POLLUTANTS IN DALLAS, TEXAS: IMPLICATIONS FOR TIME-SERIES EPIDEMIOLOGIC STUDIES

Abstract

A large scale time-series study of acute health effects and ambient air pollution in Dallas, Texas in currently underway. Air pollution measurement error due to spatial variability can impact the results of such studies; therefore, an assessment of the spatial variability of ambient air pollutants in this region is of interest. The spatio-temporal variability of daily metrics of 12 ambient air pollutants was analyzed for the 6-year period of 2003-2008. Semivariograms were constructed and spatial variability of pollutants was compared to results of an air pollution measurement error analysis in Atlanta, Georgia. The dual urban centers of Dallas and Fort Worth were observed to impact the distribution of pollutant concentrations, suggesting that an anisotropic model of pollutant concentrations may be most appropriate for characterization of ambient pollutants in this region. Semivariogram sensitivity analysis, however, demonstrated that spatial autocorrelation could be modeled isotropically for this region. SO_2 was found to be highly variable over space, likely due to presence of several point emission sources within the study area. On the basis of this initial assessment of spatial variability, the impact of exposure measurement error due to spatial variability on health risk estimates in the time-series study is estimated to be slightly less than impacts found in Atlanta, due to a greater degree of spatial autocorrelation between monitors observed for all pollutants studied.

6.1 Introduction

Time-series epidemiologic studies looking at ambient air pollution levels rely on air pollutant measurements from fixed outdoor monitoring sites to assess short term changes in air quality against a metric for an acute health outcome. Because such studies require long time-series to gain significant statistical power, the availability of continuous daily air pollutant measurements is essential. Further, time-series studies often utilize a single measure to represent population exposure to an ambient air pollutant on a given day; thus, the level to which this measure characterizes ambient levels over a broader spatial area is central to the interpretation of time-series studies. Pollutant concentrations can vary widely over space even within a single urban area (Hewitt 1991) and as a result, the choice of metric used in time-series studies has been shown to impact the health association observed (Chen et al. 2007; Sarnat et al. 2010; Wilson et al. 2007). Additionally, health associations have been shown to differ between cities (Samet et al. 2000). A large scale time-series study of emergency department visits and ambient air pollutants in Dallas, Texas is currently underway; thus, a thorough examination of the air pollutant distribution as it will impact the health study is of interest. This analysis assesses the spatio-temporal distribution of 12 ambient air pollutants in the Dallas metropolitan area and discusses results in the context of a planned time-series health study.

6.2 Methods

6.2.1 Air Pollutant Data

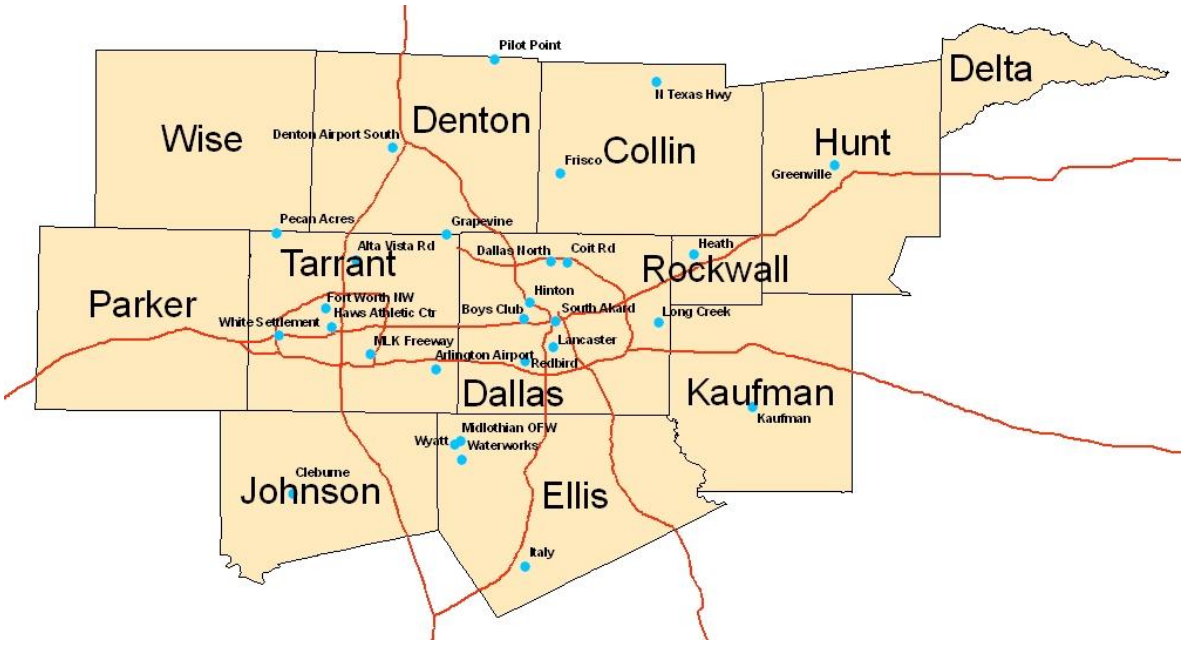
Data for the 2003-2008 time period were collected for 11 counties in the Dallas metropolitan area: Collin, Dallas, Delta, Denton, Ellis, Hunt, Johnson, Kaufman, Parker, Rockwall, Tarrant, and Wise. Daily and/or hourly metrics were collected for 12 ambient air pollutants: NO_2, NO_x, O_3, SO_2, CO, PM_{10} mass, $PM_{2.5}$ mass, and $PM_{2.5}$ components

sulfate (SO_4), nitrate (NO_3), ammonium (NH_4), elemental carbon (EC) and organic carbon (OC). A map of the study area with monitor locations is shown in Figure 6.1. Air pollutant measures were utilized from 2 data sources: the US EPA's Air Quality System (AQS), including the Texas Commission on Environmental Quality (TCEQ) Continuous Ambient Monitoring Stations (CAMS) and the Chemical Speciation Network (CSN) for $PM_{2.5}$ component measurements and the Southeastern Aerosol Research and Characterization Study (SEARCH) network which collected daily measures of $PM_{2.5}$ and its components for 2006-2007 only.

Hourly gas data were used to compute daily 1-hour maxima (NO_2, NO_x, CO, SO_2) and daily 8-hour maxima (O_3). NO_x measurement includes NO, NO_2 as well as other oxides of nitrogen. For $PM_{2.5}$ mass, data from different measurement methods were used at some monitoring sites. Both Federal Reference Method (FRM) instruments, which provide 24-hr average filter-based measurements, and Tapered Element Oscillating Microbalance (TEOM) instruments, which provide 1-hr semi-continuous data, were utilized. The TEOM analyzers were operated to minimize loss of semi-volatiles (Edgerton et al. 2006) and discrepancy between measurements is expected to have little impact on this analysis. For $PM_{2.5}$ components, data were obtained from particle composition monitors (PCM) providing 24-hr filter samples and ions (SO_4, NO_3, NH_4) were detected by chromatographic analysis. Thermal optical transmittance method was utilized for measurement of carbonaceous species (EC and OC).

In this study, Dallas measures are juxtaposed against comparable measures from a previous air pollution epidemiologic study of Atlanta, Georgia (Metzger et al. 2004). The monitoring network and measurement methods utilized in the Atlanta air quality dataset are described in detail in previous work (Goldman et al. 2010). Differences in measurement methods between the Atlanta and Dallas central monitoring sites are highlighted here. The Hinton Street monitor, which is located 8 miles west of downtown Dallas, is designated the central monitoring site for the Dallas study and the Jefferson

Street monitor, located 3 miles northwest of downtown Atlanta, is considered the central monitoring site for Atlanta. At Hinton Street, AQS measures were utilized for all pollutants, while at Jefferson Street, SEARCH data were used for all measures. Several measurement differences between the networks are noted. SEARCH monitors independently measure NO and NO_2, and therefore, do not include additional oxides of nitrogen in their measurement as AQS does. For SO_2 measurement, less SO_2 is lost by water condensation in the SEARCH sampling system than in the AQS measurement. For carbonaceous species, SEARCH uses the thermal optical reflectance method for differentiating EC and OC, whereas AQS uses the thermal optical transmittance method. The EC-OC split differs between these methods (Chow et al. 2001); however, the measures are highly correlated and as a result, the impact of this difference on this analysis is expected to be minimal (Solomon et al. 2003a).



Site Label	Pollutants Measured
Alta Vista Rd	O_3
Arlington Airport	NO_2/NO_x, O_3, CO
Boys Club	PM_{10}, $PM_{2.5}$
Cleburne	O_3
Coit Rd	PM_{10}, $PM_{2.5}$
Dallas North	NO_2/NO_x, O_3
Denton Airport South	NO_2/NO_x, O_3
Fort Worth NW	NO_2/NO_x, O_3, CO, $PM_{2.5}$
Frisco	O_3
Grapevine	NO_2/NO_x, O_3
Greenville	NO_2/NO_x, O_3
Haws Athletic Ctr	O_3, $PM_{2.5}$
Heath	O_3
Hinton	NO_2/NO_x, O_3, SO_2, CO, $PM_{2.5}$ + components
Italy	NO_2/NO_x, O_3, SO_2
Kaufman	NO_2/NO_x, O_3, SO_2, $PM_{2.5}$ + components
Lancaster	$PM_{2.5}$
Long Creek	O_3
Midlothian OFW	NO_2/NO_x, O_3, SO_2, PM_{10}, $PM_{2.5}$ + components
MLK Freeway	$PM_{2.5}$
N Texas Hwy	O_3
Pecan Acres	O_3
Pilot Point	O_3
Red Bird	NO_2/NO_x, O_3, $PM_{2.5}$
South Akard	PM_{10}, $PM_{2.5}$ + components
Waterworks	NO_2/NO_x, O_3, SO2, PM_{10}, $PM_{2.5}$
White Settlement	$PM_{2.5}$, PM_{10}
Wyatt	NO_2/NO_x, SO_2, PM_{10}

Figure 6.1. Map of monitoring site locations and key of pollutants measured. County lines and expressways are shown.

6.2.2 Spatial Assessment of Concentrations

To assess the spatial structure of concentrations for each pollutant, the mean and standard deviation of measurement time-series at each monitoring site were plotted spatially. This assessment was utilized to estimate the annual mean and standard deviation of concentrations over space for the 12-county Dallas-Fort Worth region. Modeled trends are estimated for the annual mean and annual standard deviation for all 12 pollutants.

6.2.3 Spatial Autocorrelation Assessment

The correlation between the central monitor time-series and all other monitoring sites was calculated and plotted for each pollutant. All monitor pairs with significant coinciding measurements (>365) were included in the analysis. To assess if the correlation structure was consistent at different points in space, two additional figures were constructed for $PM_{2.5}$ mass, in which correlation between all monitors and the downtown Fort Worth monitor and correlations between all monitors and the Denton Airport South monitor were plotted.

The geostatistical tool of the semivariogram provides information on the spatial autocorrelation of data and has been used previously to assess air pollutant measurement error due to spatial variability for use in time-series epidemiologic studies of ambient air pollution (Goldman et al. 2010; Goldman et al. 2011). Semivariograms were computed on normalized logged concentrations, as follows.

$$\chi_i^* = \frac{\ln C_i - \mu_{\ln C}}{\sigma_{\ln C}} \sim \mathrm{N}(0,1) \qquad (6.1)$$

Here, χ_i^* is the normalized log concentration on day *i* and $\mu_{\ln C}$ and $\sigma_{\ln C}$ are the mean and standard deviation, respectively, of the log concentrations over all days *i*; thus, the mean and standard deviation of χ_i^* are 0 and 1, respectively. A modified semivariogram was then computed on this normalized value. Here, the semivariogram is computed as the

difference between normalized observations (χ_k and χ_l) at two locations (*k* and *l*) located a distance *h* apart normalized by the temporal variance (variation over the time-series of observations) of the average of two normalized observations to yield a scaled semivariance, γ'. This scaled semivariance (i.e. the semivariance of normalized values) is related to the Pearson correlation coefficient (R) between normalized observations from two monitors as follows (Wade et al. 2006).

$$\gamma'(h) \equiv \frac{Var\left(\frac{\chi_k - \chi_l}{2}\right)}{Var\left(\frac{\chi_k + \chi_l}{2}\right)} = \frac{1 - R(h)}{1 + R(h)} \tag{6.2}$$

Thus, γ' represents the spatial semivariance scaled to a quantity indicative of the range of exposures over which health risk is being assessed. Because the value is unitless, it allows for comparison across pollutants. A scaled semivariance value of 0 corresponds to perfectly correlated observations (R = 1) and a value of 1 corresponds to perfectly uncorrelated observations (R = 0).

Correlations between observations from all pairs of monitors measuring the same pollutant during 2003-2008 were calculated. Scaled semivariograms were constructed and modeled as a function of the distance between observations, *h*. Semivariogram nugget values were derived from a previous assessment of instrument precision error using time-series data from collocated instruments (Goldman et al. 2010). Sill values were constrained to 1, i.e. R = 0, for all pollutants, assuming a distance at which any two monitors would no longer exhibit any observable correlation. Least squares regression was applied to determine the range for each pollutant.

To test the sensitivity of the semivariogram model to specific monitor time-series, additional semivariogram models were computed for NO_x and $PM_{2.5}$ mass, in which a semivariance model was regressed for the raw semivariances specific to the downtown Dallas monitor, the downtown Fort Worth monitor and the Denton Airport South

monitor, located more than 45 km from both Dallas and Fort Worth city centers. Differences between the three models are assessed. Additionally, semivariogram models were constructed for a 5% increase and 5% decrease of inter-monitor correlations, in order to assess the sensitivity of the curves to Pearson correlation coefficient values.

6.3 Results

6.3.1 Intercity Comparison

Dallas and Atlanta study areas were compared to assess differences that may affect air quality and its associated health effects (Table 6.1). The spatial area included in the Dallas study is larger than that in Atlanta and the study population is roughly one million persons greater, with a total study population of slightly greater than 6 million based on 2008 data. The average Dallas population density of 272 persons/km^2 is lower than the 331 persons/km^2 in Atlanta.

Population density by census tract was plotted for Atlanta and Dallas to assess the spatial distribution of population in the two cities (Figure 6.2). For Atlanta, the regions of greatest population are located near the urban center and density tends to decrease uniformly in all directions as distance from downtown increases. In Dallas, the population trends observed are less symmetric. The greatest populations are observed near the Fort Worth and Dallas urban centers, with high population densities along the east-west line between the two cities. Density decreases to the north and south, as well as to the east of Dallas and to the west of Fort Worth. This illustrates that population trends in the Dallas-Fort Worth region follow more of an ellipsoid pattern, with population density decreasing more rapidly in the north-south direction than in the east-west direction with respect to the regional center. Given that many ambient air pollutants, particularly traffic-related species, tend to correlate with population density, this

observation is informative for assessing the spatial distribution of ambient concentrations and creates an interesting challenge from a geostatistical modeling point of view.

Table 6.1. Study area comparison between Atlanta, GA and Dallas, TX.

City	Number of Counties	Area (km^2)	Population (2008)	Population Density (pop/km^2)
Dallas	12	23,120	6,300,006	272
Atlanta	20	15,787	5,218,150	331

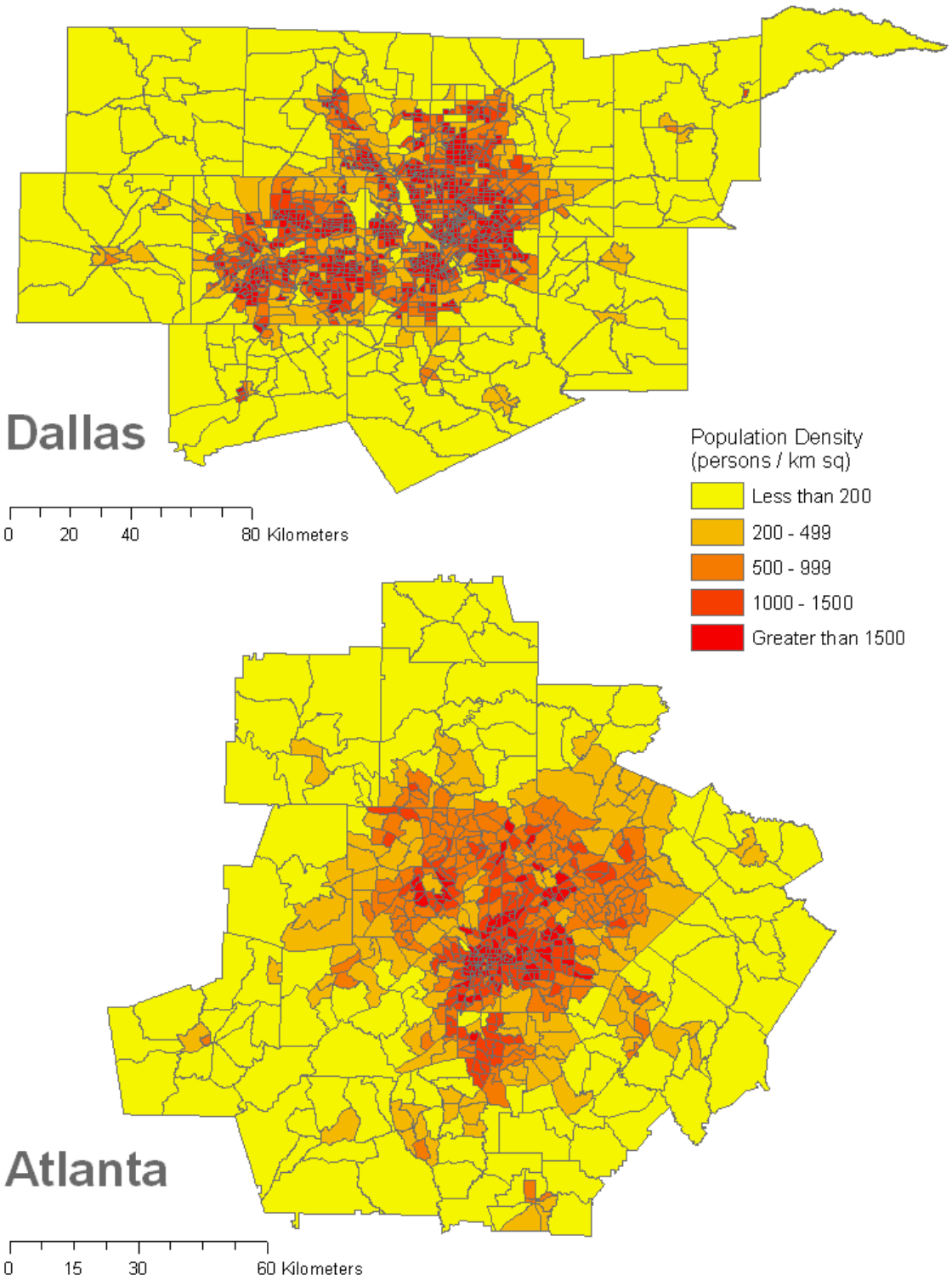


Figure 6.2. Population density (persons/km^2) for the Atlanta and Dallas-Fort Worth regions.

6.3.2 Air Pollutant Data

Daily metrics for the 12 ambient air pollutants were collected for all available monitoring sites in the Dallas metropolitan area measuring during the study period of 2003-2008. Data from several monitors operated for the duration of the study period with minimal missing data were available for each pollutant through the AQS network, ranging from 2 complete time-series available for $PM_{2.5}$ components and 14 complete time-series available for O_3. Other sites with multi-year time-series over part of the study period were available for several pollutants. Those with data time-series of significant length are of use for the spatial variability assessment. There were differences in the frequency at which each pollutant was measured. All gases were measured on an hourly basis at all sites and daily 1-hour maxima were computed. Twenty-four hour samples of PM_{10} mass were measured at a frequency of six days at all sites. For $PM_{2.5}$ mass, sites with a TEOM provided hourly measurements, while FRM sites provided 24-hour samples and frequencies that varied by site. Three sites measured daily for the duration of the study period, while the other sites measured at a frequency of three or six days. $PM_{2.5}$ components were measured at a frequency of three days at both sites that provided complete time-series and every six days at three other sites with component measurements available. With respect to missing values, data completeness ranged from 90% for EC/OC to 99% for gaseous species. Table 6.2 lists the number of sites and data completeness range for each pollutant.

Table 6.2. Measurement data completeness. Number of sites measuring for entire 2003-2008 period with the number of additional sites providing data for a portion of the time period listed parenthetically. Minimum to maximum percent completeness for monitors measuring for the entire 2003-2008 period is listed.

Pollutant	Number of Sites	Percent Completeness
NO_2	10 (3)	91% - 99%
NO_x	10 (3)	91% - 99%
O_3	14 (6)	95% - 99%
SO_2	3 (3)	94% - 98%
CO	3	92% - 99%
PM_{10}	3 (4)	82% - 96%
$PM_{2.5}$	10 (9)	86% - 99%
$PM_{2.5}$-SO_4	2 (3)	82% - 91%
$PM_{2.5}$-NO_3	2 (3)	81% - 91%
$PM_{2.5}$-NH_4	2 (3)	82% - 91%
$PM_{2.5}$-EC	2 (3)	82% - 90%
$PM_{2.5}$-OC	2 (3)	82% - 90%

The geometric mean and standard deviation of the 6-year pollutant time-series at the Dallas central monitor, Hinton Street, are compared with those of the Atlanta central monitor, Jefferson Street (Table 6.3). Concentrations at the central monitors were largely similar for Atlanta and Dallas with some notable differences. Ozone, PM_{10} mass, $PM_{2.5}$ mass and the nitrate component of $PM_{2.5}$ exhibited similar levels between cities. Traffic-related pollutants (NO_2, NO_x, CO and $PM_{2.5}$-EC) were greater in Atlanta compared to levels in Dallas. While both cities are expected to have substantial impacts of traffic emissions, elevated levels of these pollutants in Atlanta may be a reflection of meteorological differences between the two cities. SO_2 concentrations are tenfold higher at the Atlanta downtown monitor, likely due to coal-fired power plants located in proximity to metropolitan Atlanta (Lowe 2007); $PM_{2.5}$-SO_4 levels are similarly elevated. The organic carbon component of $PM_{2.5}$ is also greater in Atlanta than Dallas, which may be the result of elevated levels of secondary organic carbon in the southeast due to high biogenic emissions (Lee et al. 2010; Pachon et al. 2010).

It is also of interest to look at the relative proportions of particulate matter components between the two cities. Figure 6.3 illustrates average $PM_{2.5}$ speciation data

for 2008. Notably, organic carbon is a more dominant contributor to the overall $PM_{2.5}$ mass for Atlanta (35% of $PM_{2.5}$ mass) compared to Dallas (26% of $PM_{2.5}$ mass). Sulfate, nitrate and ammonium were larger fractions of $PM_{2.5}$ in Dallas than Atlanta.

Table 6.3. Geometric mean and standard deviation of air pollutant measurements for the Dallas Hinton Street monitor (2003-2008) and Atlanta Jefferson Street monitor (1999-2004)

	Atlanta		Dallas	
	μ_g	σ_g	μ_g	σ_g
1-hr max NO_2	38.1 ppb	1.51	24.4 ppb	2.31
1-hr max NO_x	86.8 ppb	2.23	40.9 ppb	3.22
8-hr max O_3	35.8 ppb	1.92	35.0 ppb	1.66
1-hr max SO_2	11.4 ppb	2.58	1.88 ppb	2.06
1-hr max CO	0.89 ppm	2.11	0.34 ppm	6.04
24-hr PM_{10}	23.8 $\mu g/m^3$	1.55	23.1 $\mu g/m^3$	1.46
24-hr $PM_{2.5}$	15.2 $\mu g/m^3$	1.62	10.4 $\mu g/m^3$	1.59
24-hr $PM_{2.5}$-SO_4	3.85 $\mu g/m^3$	1.95	2.56 $\mu g/m^3$	2.00
24-hr $PM_{2.5}$-NO_3	0.74 $\mu g/m^3$	2.13	0.68 $\mu g/m^3$	2.29
24-hr $PM_{2.5}$-NH_4	2.02 $\mu g/m^3$	1.79	1.01 $\mu g/m^3$	2.10
24-hr $PM_{2.5}$-EC	1.28 $\mu g/m^3$	1.87	0.50 $\mu g/m^3$	1.80
24-hr $PM_{2.5}$-OC	3.79 $\mu g/m^3$	1.65	2.71 $\mu g/m^3$	1.62

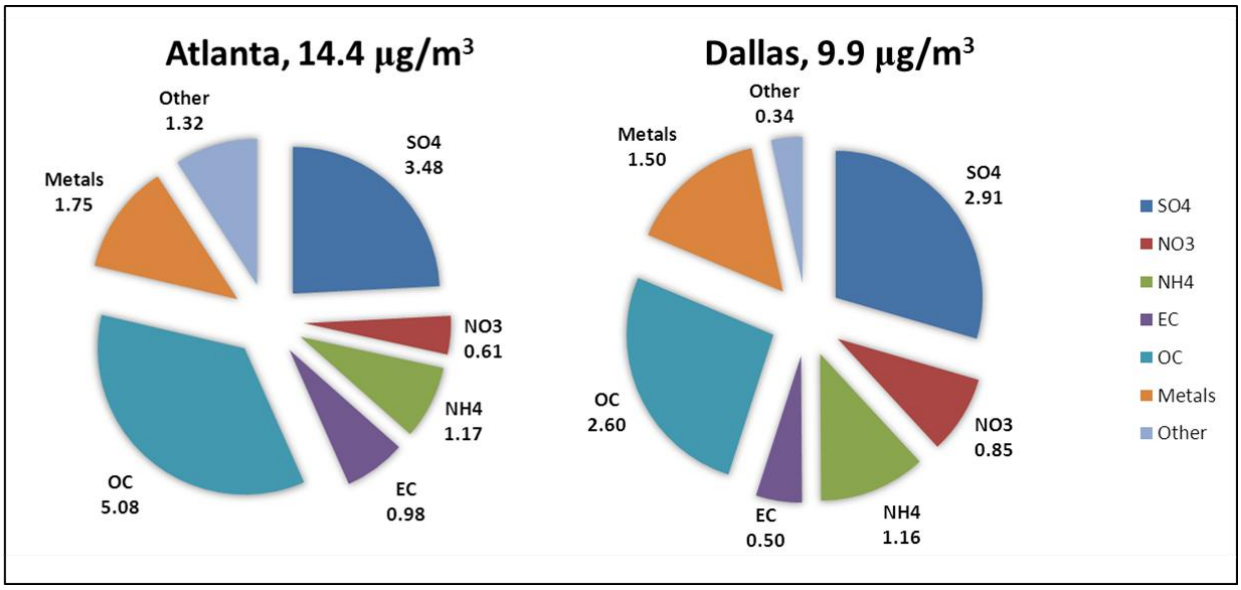


Figure 6.3. $PM_{2.5}$ Speciation for 2008 at the Atlanta and Dallas central monitoring sites: the Jefferson Street monitor in Atlanta and the Hinton Street monitor in Dallas.

6.3.3 Spatial Assessment of Concentrations

For each pollutant, mean concentrations from multiple monitors for the 2004 time-series of daily measurements were used to assess the spatial distribution of pollutant levels in Dallas. As examples of secondary, mixed source, and primary pollutants, Figures 6.4, 6.5 and 6.6 show these plots for O_3, $PM_{2.5}$ mass, and NO_x, respectively. Plots for the rest of the 12 pollutants analyzed in Dallas can be found in Appendix D.

For secondary pollutants such as O_3 (Figure 6.4), concentrations did not vary significantly over space. $PM_{2.5}$ mass concentrations, derived from both primary and secondary origin, had only a slight spatial trend over the Dallas-Fort Worth region (Figure 6.5). For primary pollutants, mean concentration decreased with distance from the two urban centers of the Dallas-Fort Worth (Figure 6.6). This suggests that, in contrast to Atlanta, the spatial distributions of primary pollutant concentrations in Dallas may not lend themselves well to an isotropic model.

The mean concentrations of NO_2 and NO_x observed at the Wyatt, Waterworks and Midlothian monitoring sites, all located southwest of Dallas in close proximity to each other, are elevated despite their significant distance from either the Dallas or Fort Worth population centers. Evidence from this analysis suggests that these sites are impacted by a local source. The monitors are included in the spatial plots, but were removed from the semivariogram analysis, as they are not considered part of the regulatory ambient concentration this analysis attempts to capture.

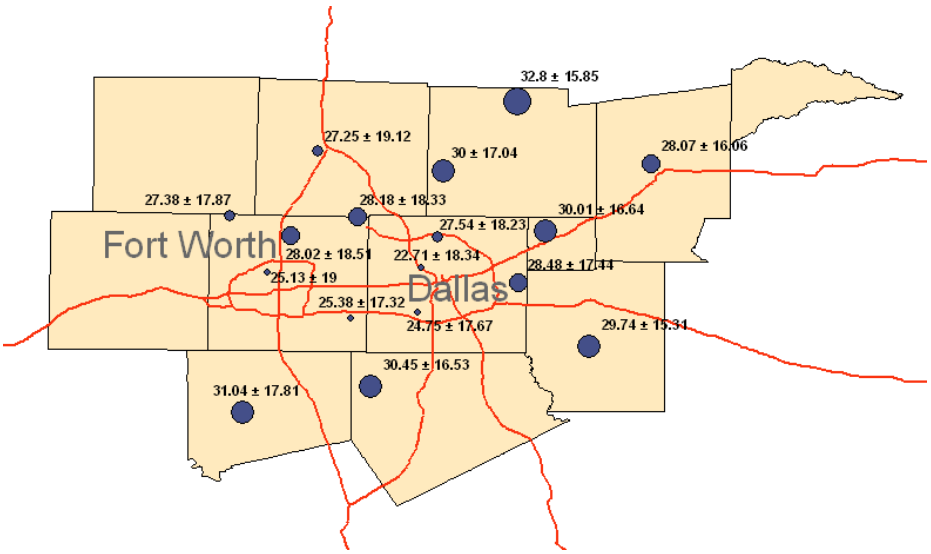


Figure 6.4. O_3 time-series mean (ppb) plotted at monitor locations with graduated symbols for the Dallas study area. Labels indicate time-series mean ± time-series standard deviation.

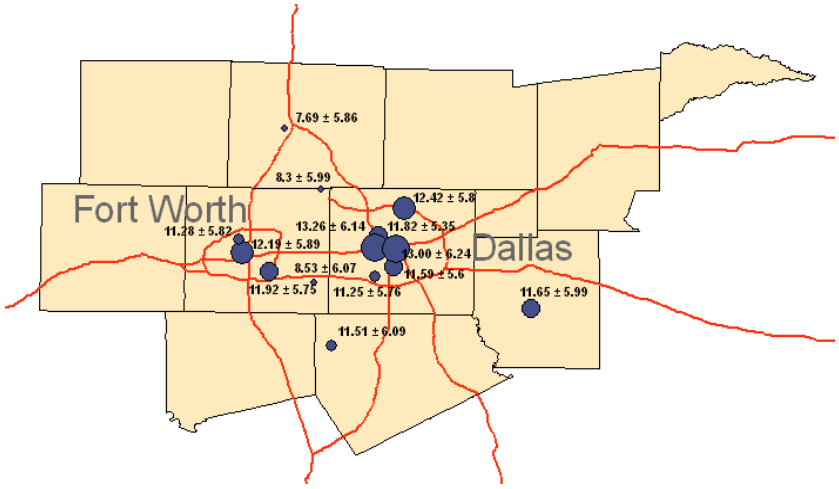


Figure 6.5. $PM_{2.5}$ mass time-series mean ($\mu g/m^3$) plotted at monitor locations with graduated symbols for the Dallas study area. Labels indicate time-series mean ± time-series standard deviation.

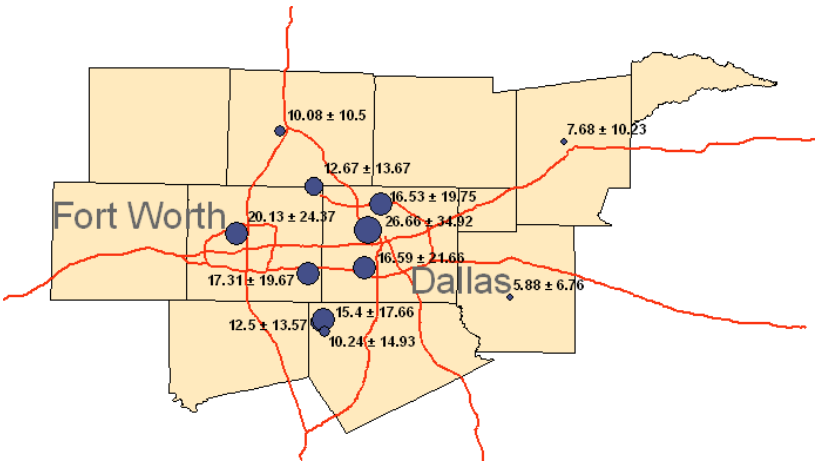


Figure 6.6. NO_x time-series mean (ppb) plotted at monitor locations with graduated symbols for the Dallas study area. Labels indicate time-series mean ± time-series standard deviation.

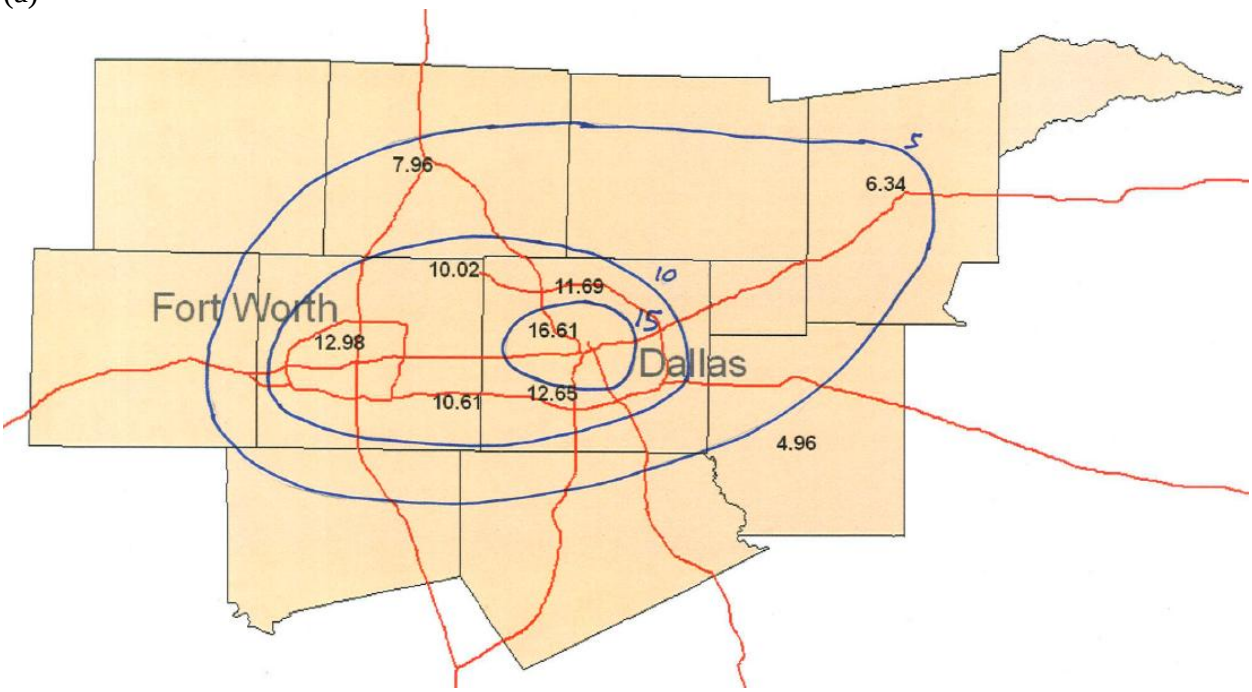
Using the spatial plots of the time-series mean and standard deviation for each monitoring site for all pollutants, isopleths of mean and standard deviation were approximated to provide a modeling framework for geostatistical assessment of annual concentration distributions (Figures 6.7-6.18). In addition to the monitor time-series mean and standard deviations, population density patterns and prior assessment of spatial distribution of pollutants in Atlanta were taken into account to estimate the spatial distribution of concentrations over the Dallas-Fort Worth region.

The NO_2 and NO_x models have the shape expected from more spatially heterogeneous pollutants and the pattern is similar to the observed population density distribution (Figure 6.1). For these pollutants, elevated annual mean and standard deviation are observed at the Dallas and Fort Worth urban centers and in between them, with values decreasing rapidly with increasing distance from this elliptical population center. Any model of the annual distribution of primary pollutants for the Dallas-Fort Worth region should account for this anisotropic phenomenon in which concentrations are expected to change more rapidly in the north-south direction than in the east-west direction, with respect to the regional center. CO and EC are also expected to follow this trend as they are traffic-related pollutions emitted primarily; however, there is insufficient

monitor coverage in rural areas to base spatial models on observations alone. Because CO and EC are highly correlated with NO_x (0.80 ± 0.12 and 0.55 ± 0.15 across corresponding sites, respectively), the urban-to-rural NO_x trend was used to approximate CO and EC trends as a function of distance from the urban centers and to estimate rural background levels. In both the Atlanta and Dallas study areas, population density drops to less than 200 persons/km^2 in census tracts far from the urban centers; thus, concentrations of traffic-related pollutants are anticipated to be drop to similar rural background levels.

Based on monitor measurements, low variance observed in the semivariogram analysis and observed trends in Atlanta, spatial trends in annual mean and standard deviation for pollutants of more secondary origin are expected to be much more gradual and some exhibit similar values over the entire study area. While concentrations of these more spatially homogeneous species are expected to eventually drop to lower background levels as distance from urban regions increases indefinitely, to the extent of the study area assessed here concentrations are more or less constant over space for some species, namely O_3, and $PM_{2.5}$ components sulfate, nitrate and ammonium. The estimated rural background mean and standard deviation of concentrations for the Dallas-Fort Worth study area are listed in Table 6.4 and the annual mean and standard deviation of measurements at the rural Yorkville site in Atlanta for 2004 are shown for comparison.

(a)



(b)

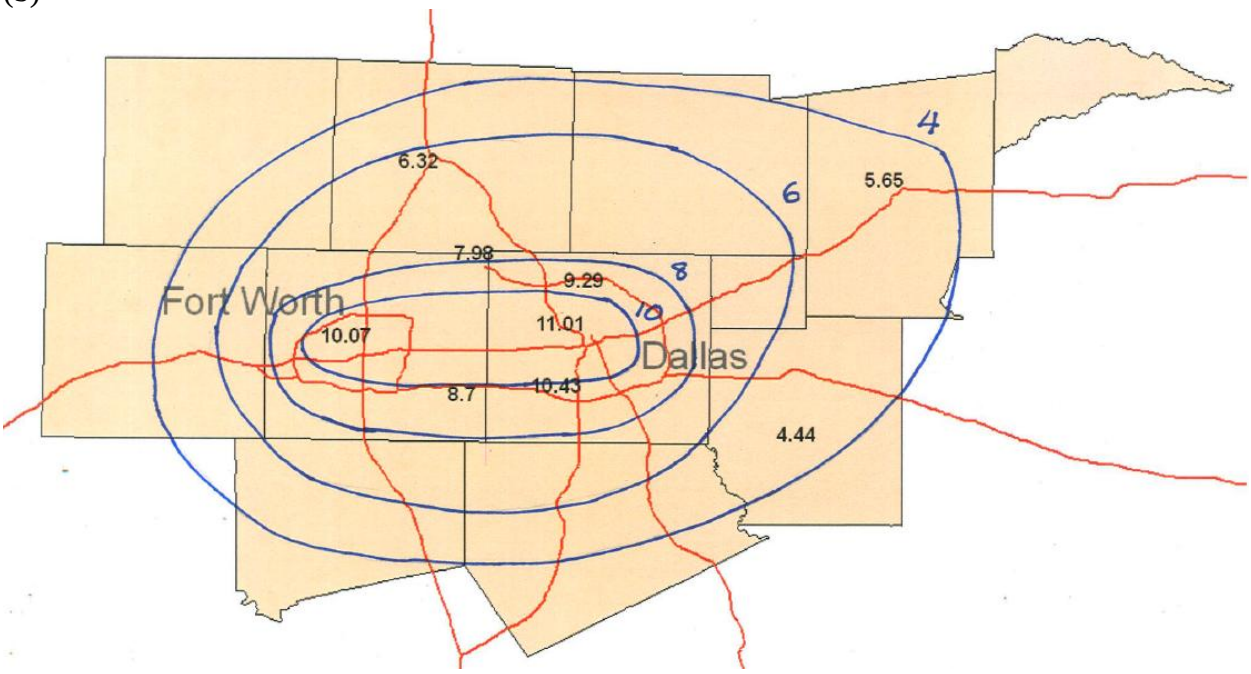
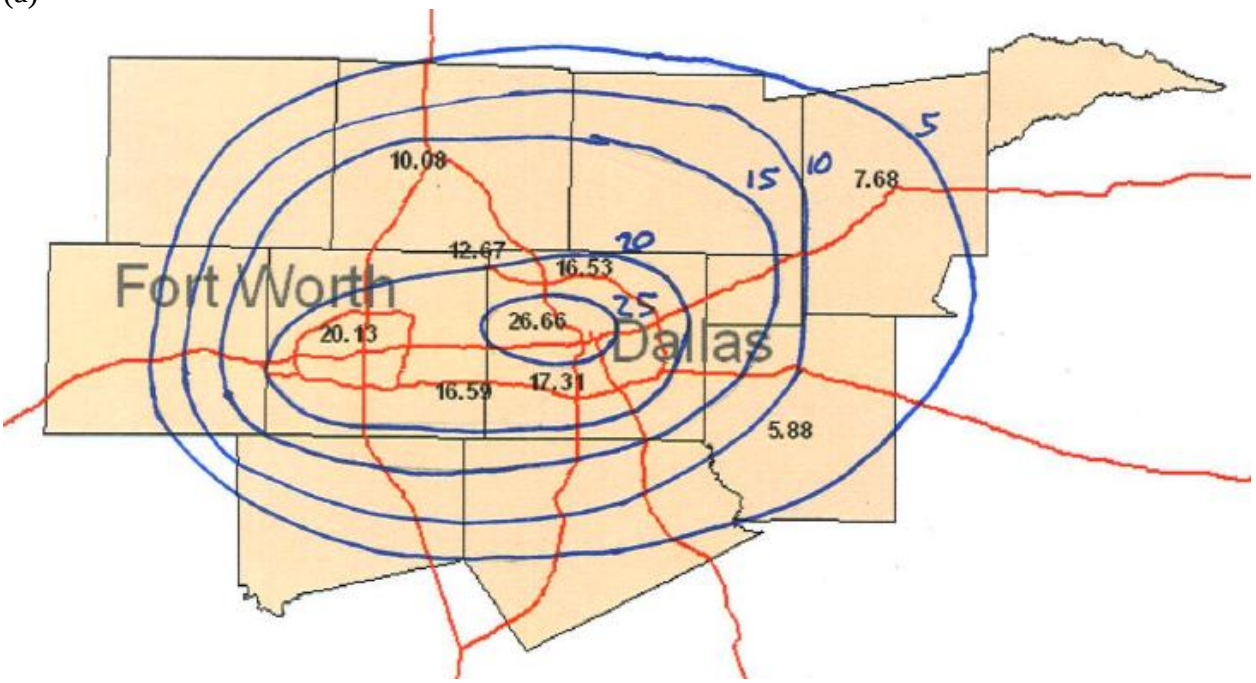


Figure 6.7. Modeled annual mean (a) and standard deviation (b) of NO_2 concentration (ppb) over the Dallas-Fort Worth region.

(a)



(b)

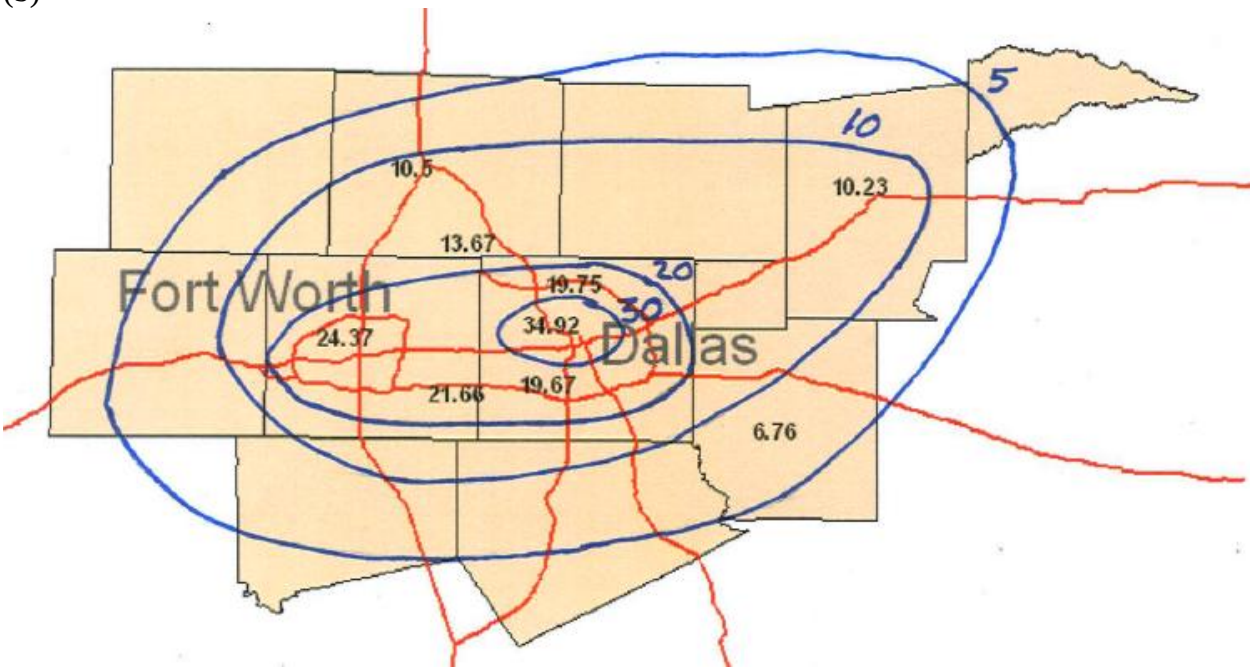
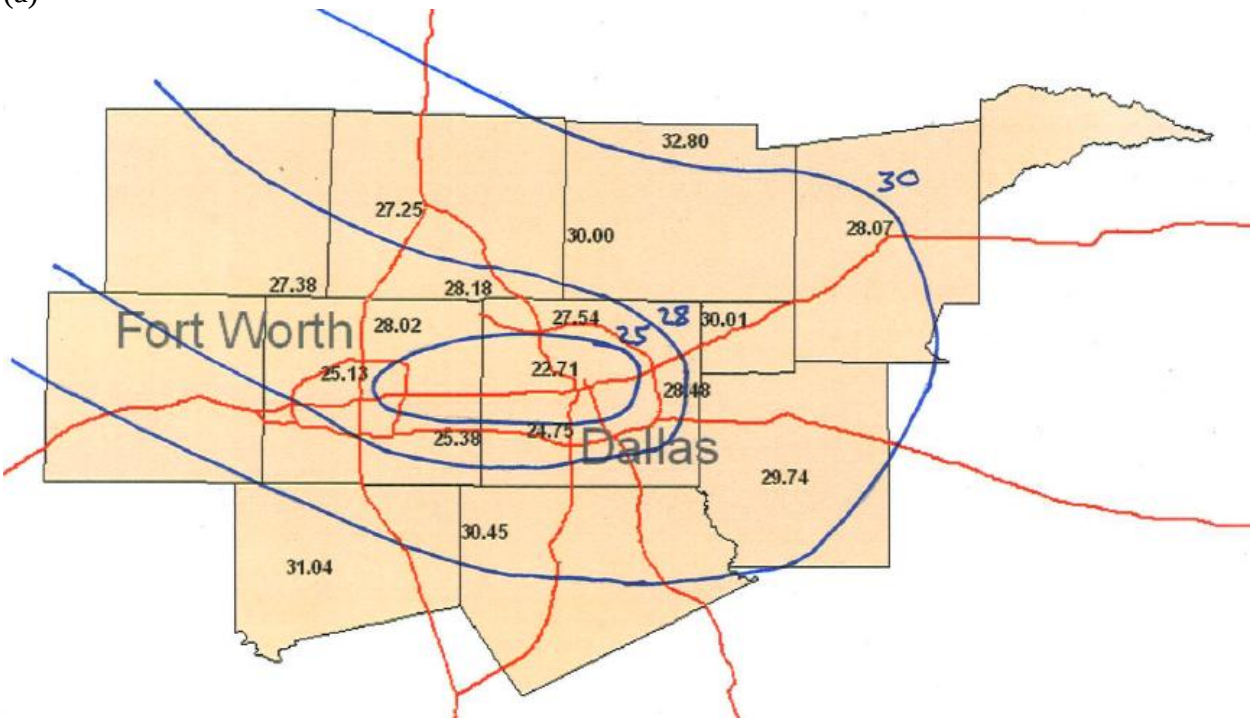


Figure 6.8. Modeled annual mean (a) and standard deviation (b) of NO_x concentration (ppb) over the Dallas-Fort Worth region.

(a)



(b)

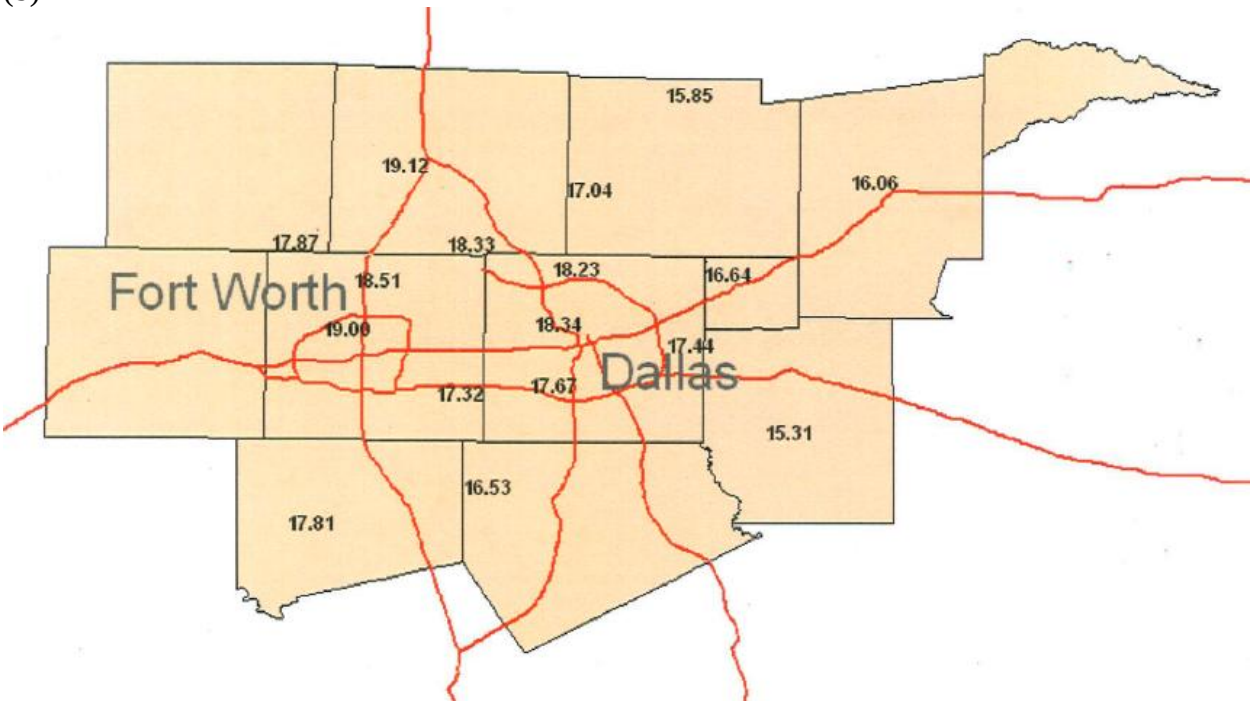
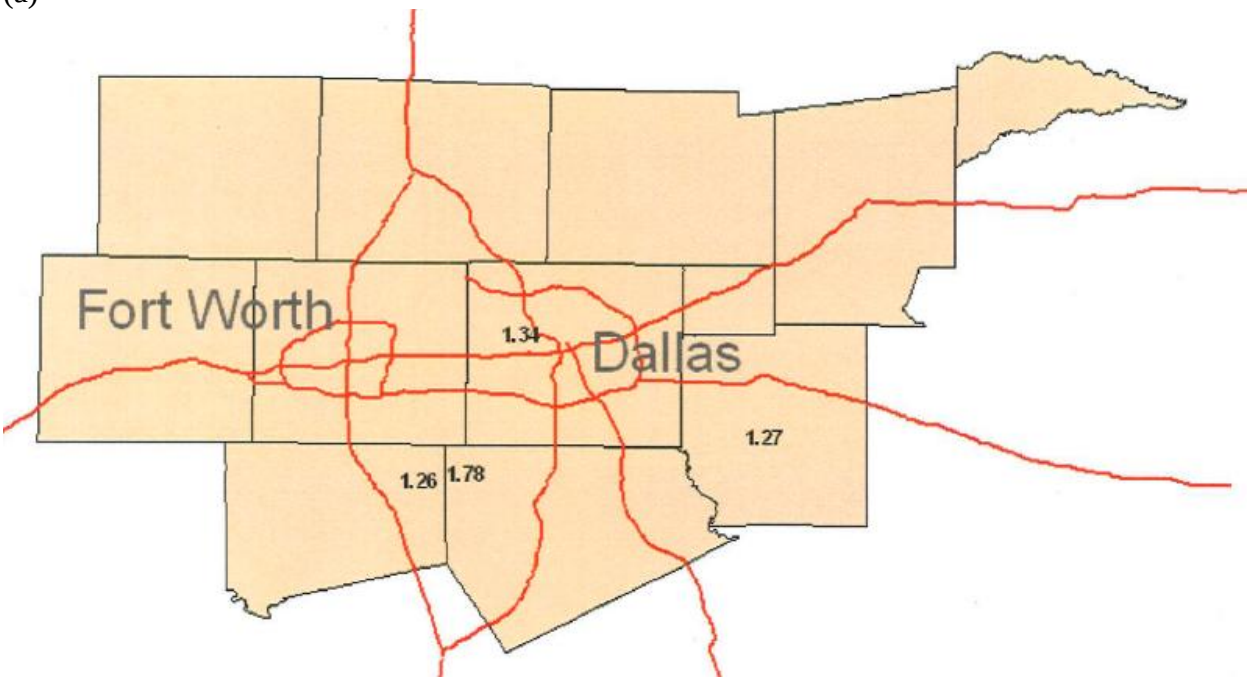


Figure 6.9. Modeled annual mean (a) and standard deviation (b) of O_3 concentration (ppb) over the Dallas-Fort Worth region.

(a)



(b)

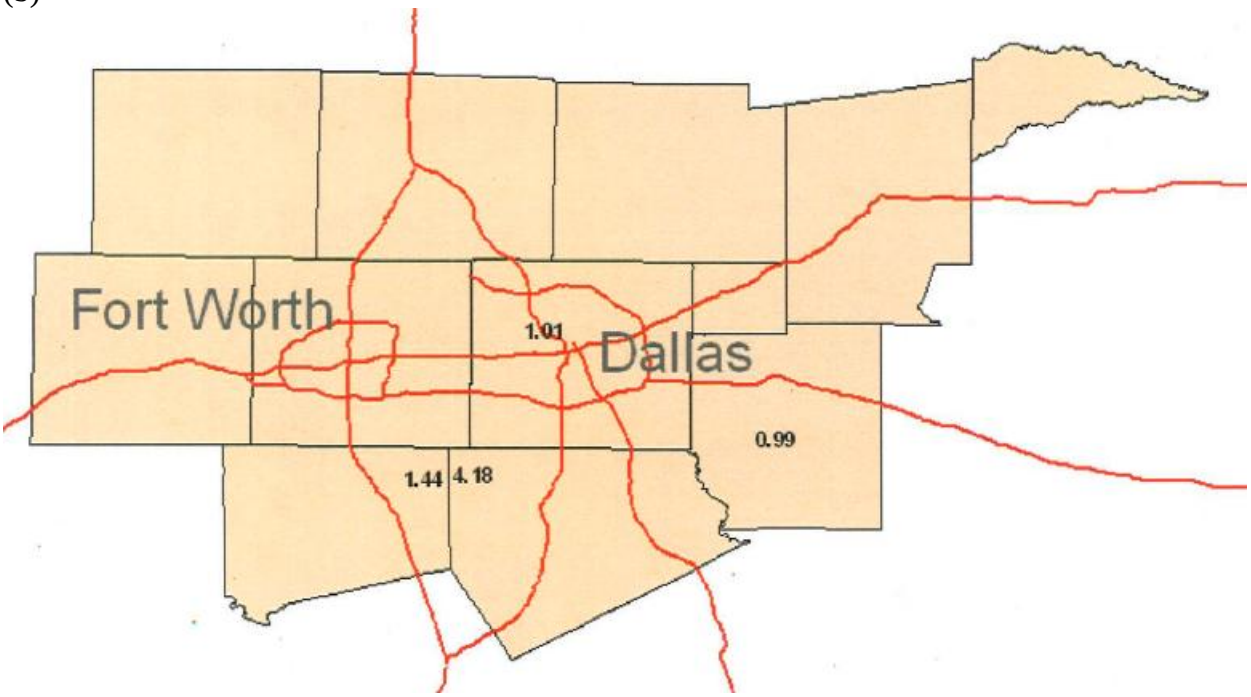
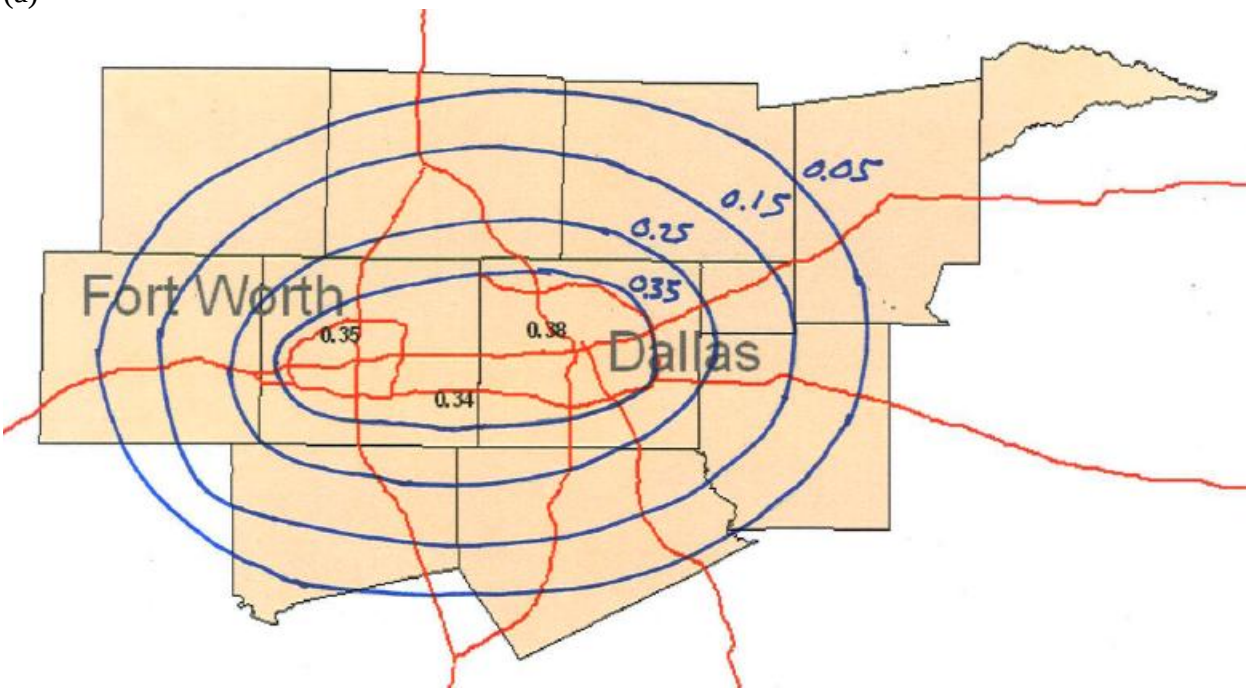


Figure 6.10. Modeled annual mean (a) and standard deviation (b) of SO_2 concentration (ppb) over the Dallas-Fort Worth region.

(a)



(b)

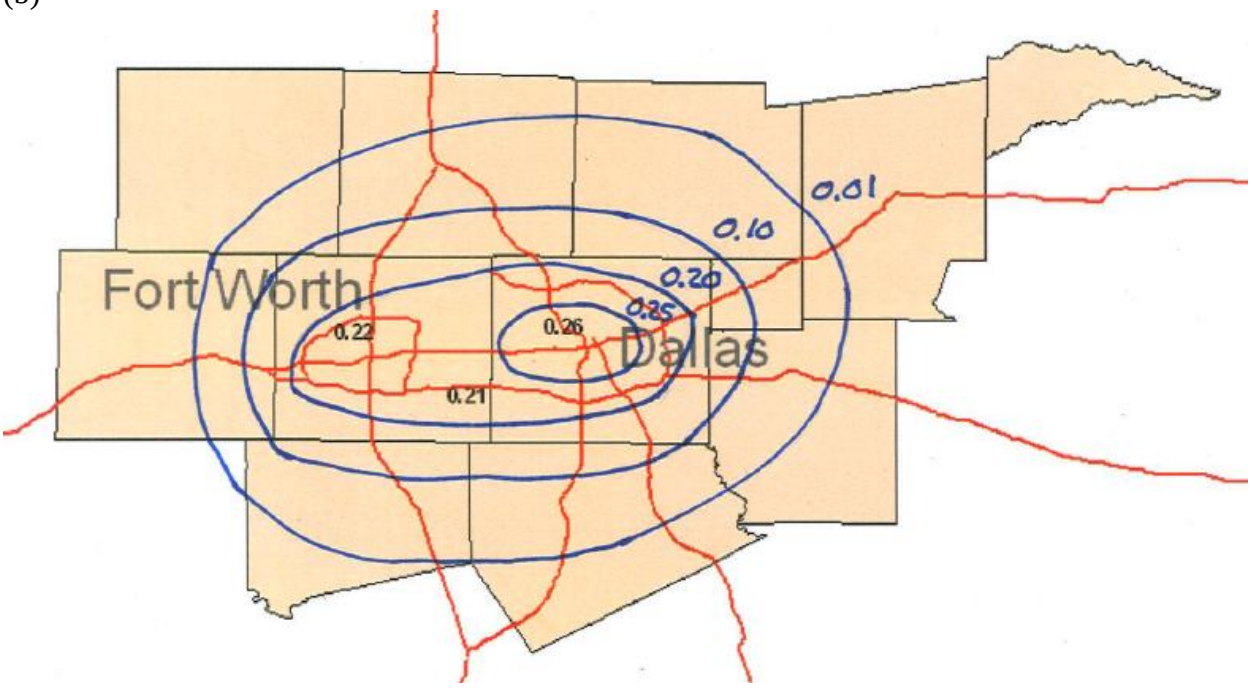
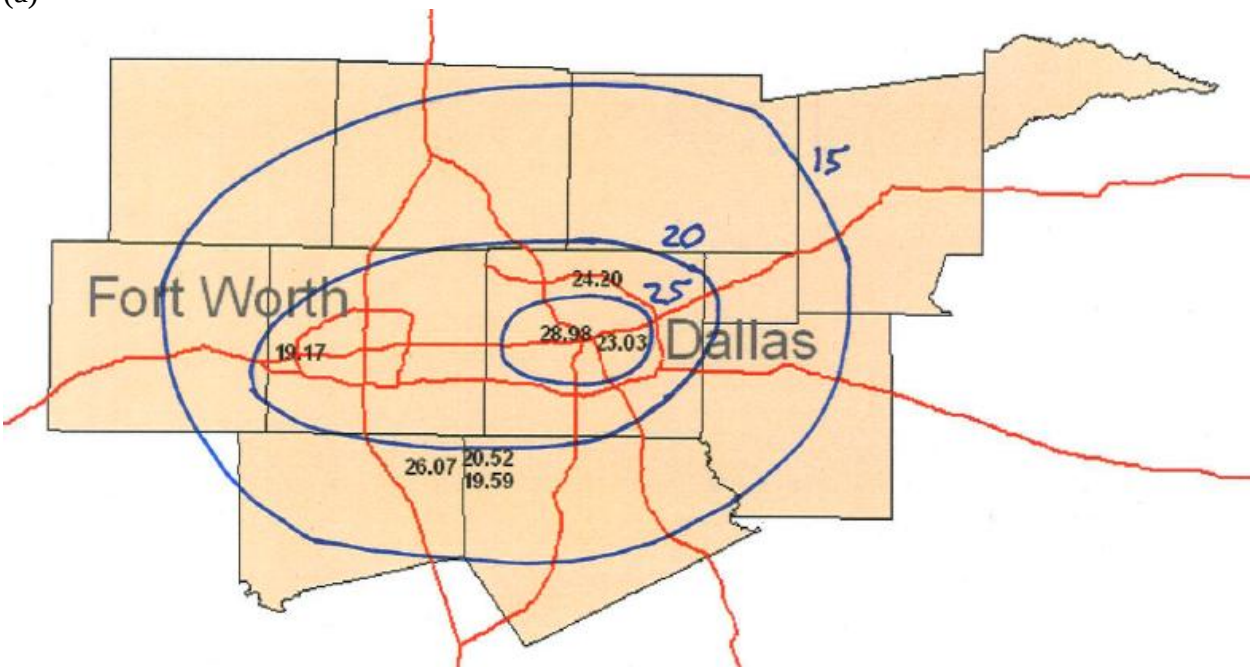


Figure 6.11. Modeled annual mean (a) and standard deviation (b) of CO concentration (ppm) over the Dallas-Fort Worth region.

(a)



(b)

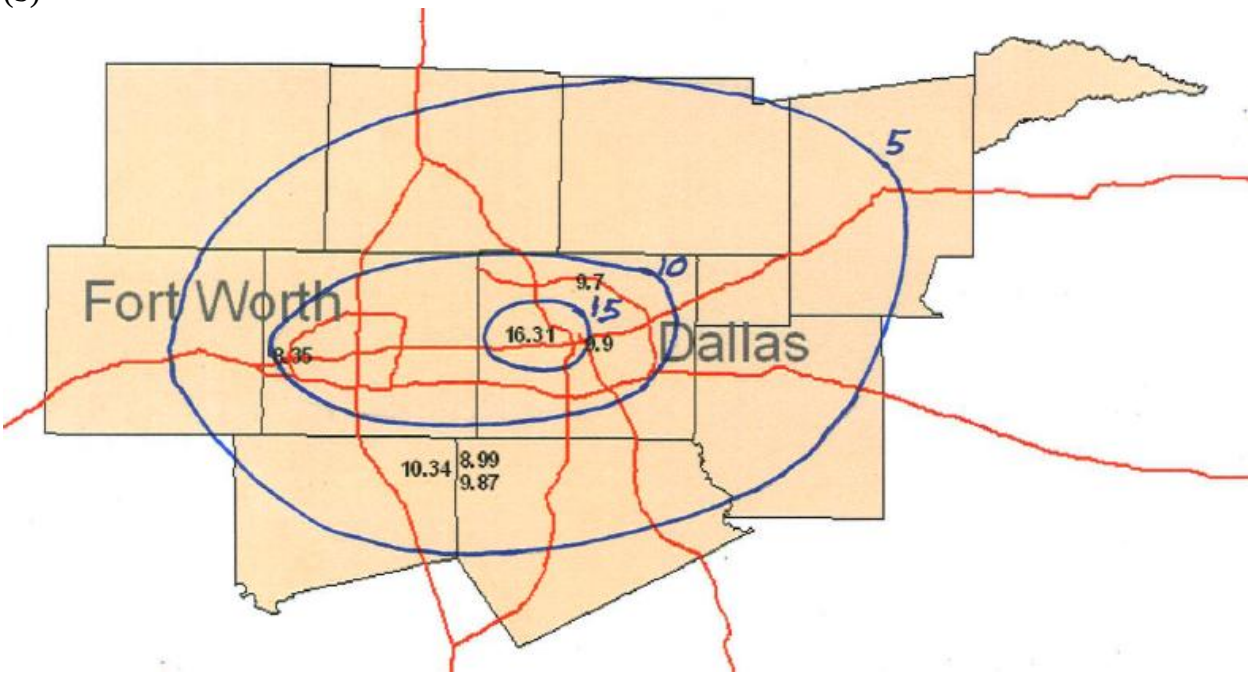
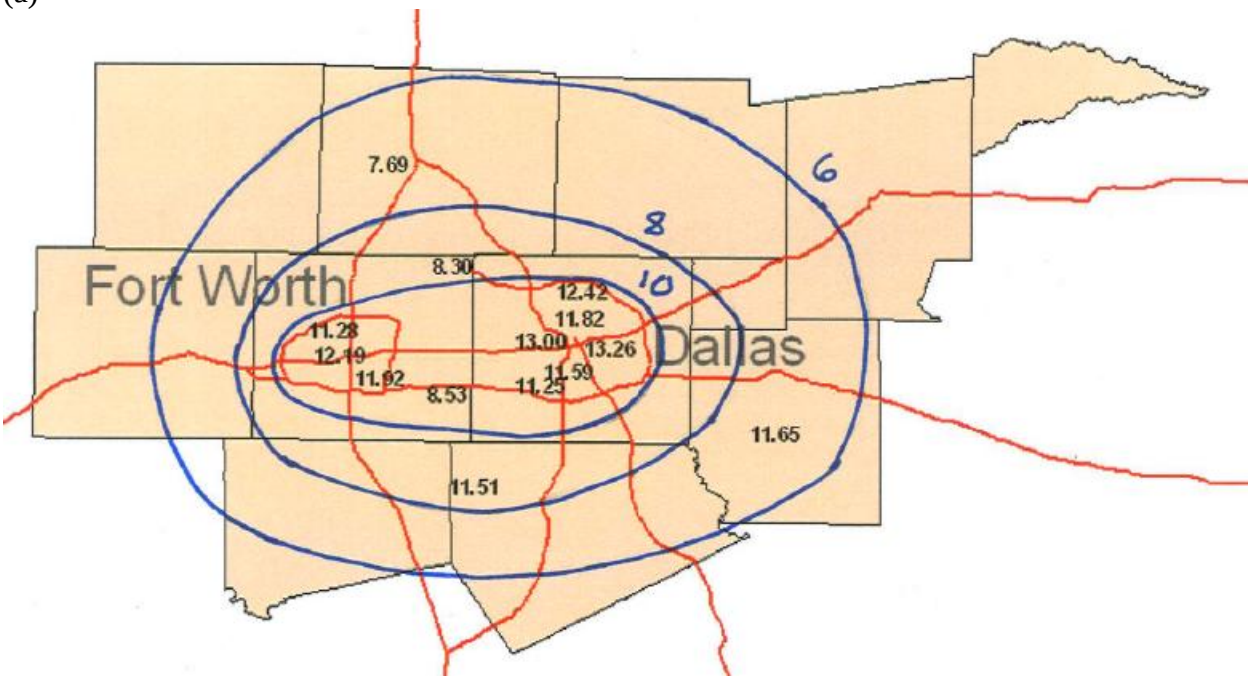


Figure 6.12. Modeled annual mean (a) and standard deviation (b) of PM_{10} mass concentration ($\mu g/m^3$) over the Dallas-Fort Worth region.

(a)



(b)

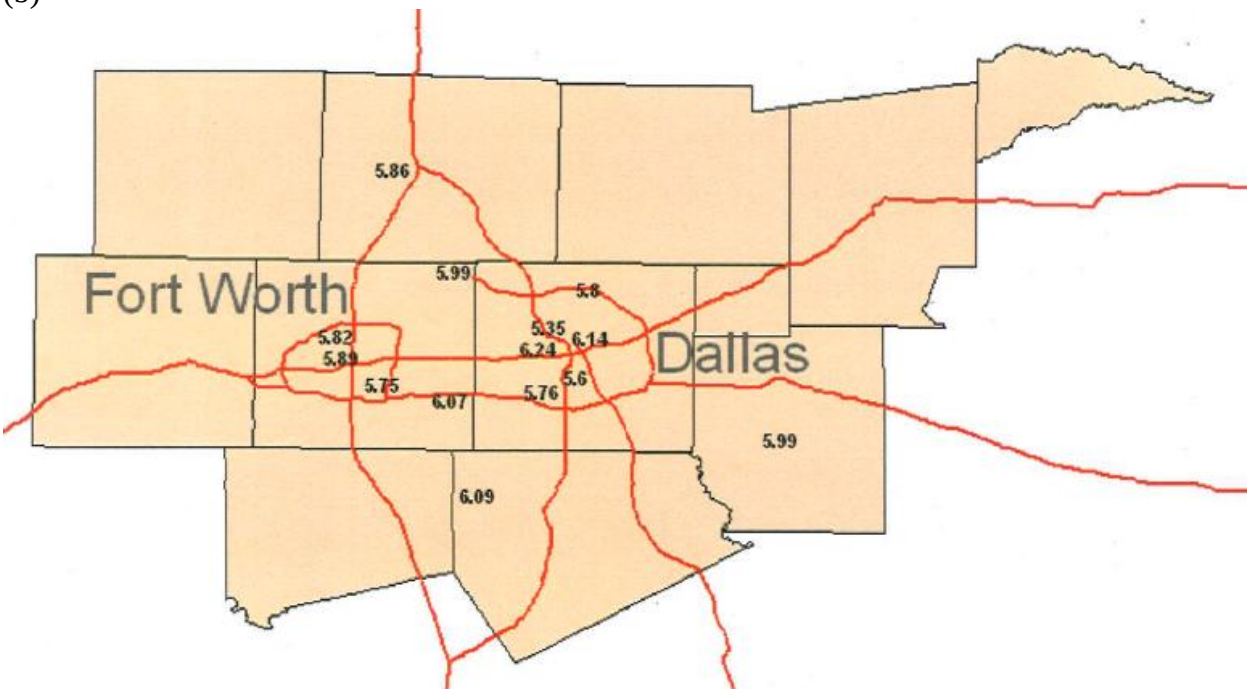
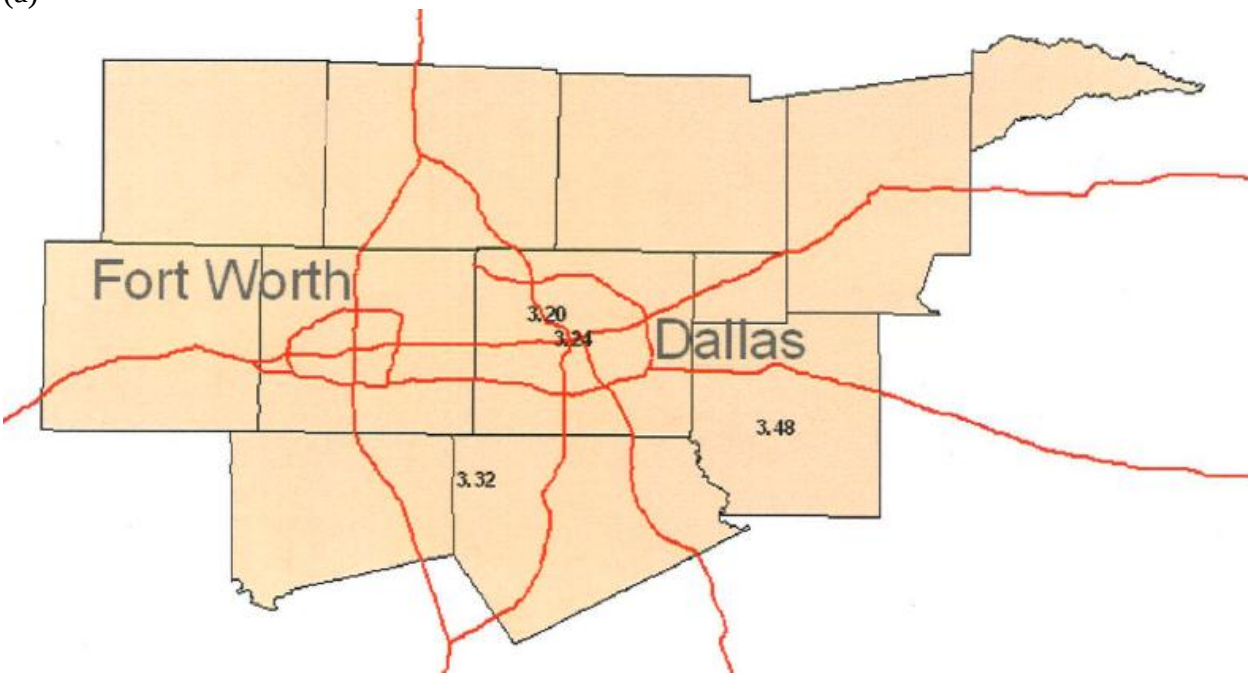


Figure 6.13. Modeled annual mean (a) and standard deviation (b) of $PM_{2.5}$ mass concentration ($\mu g/m^3$) over the Dallas-Fort Worth region.

(a)



(b)

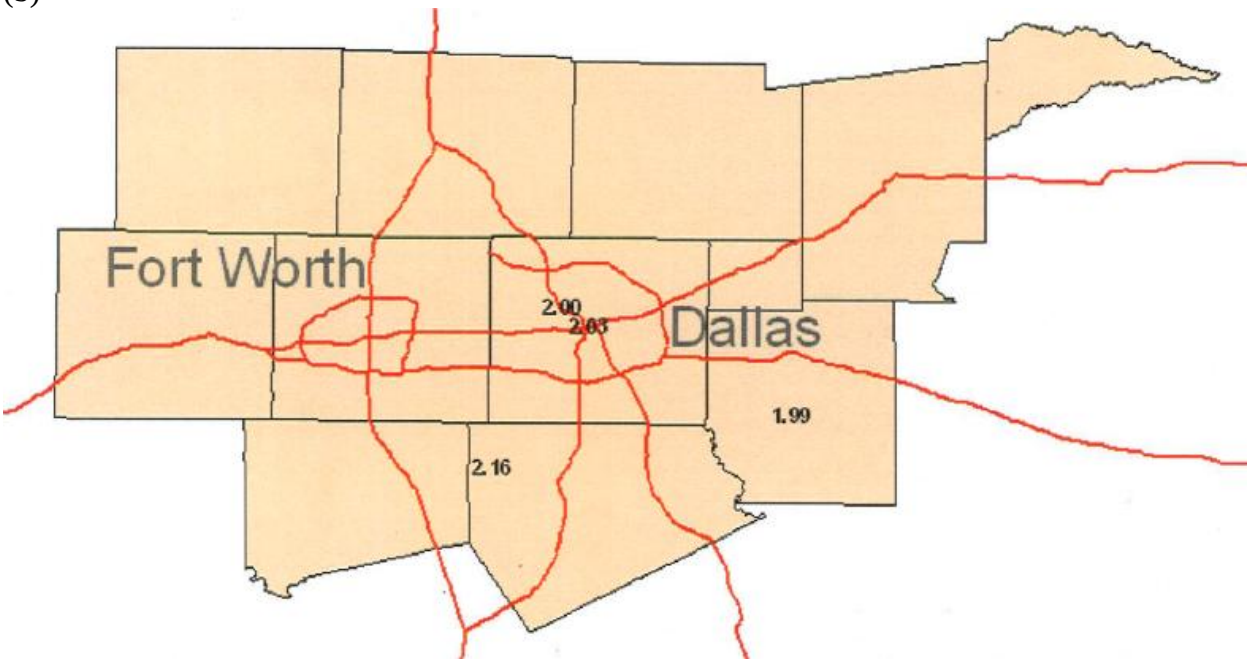
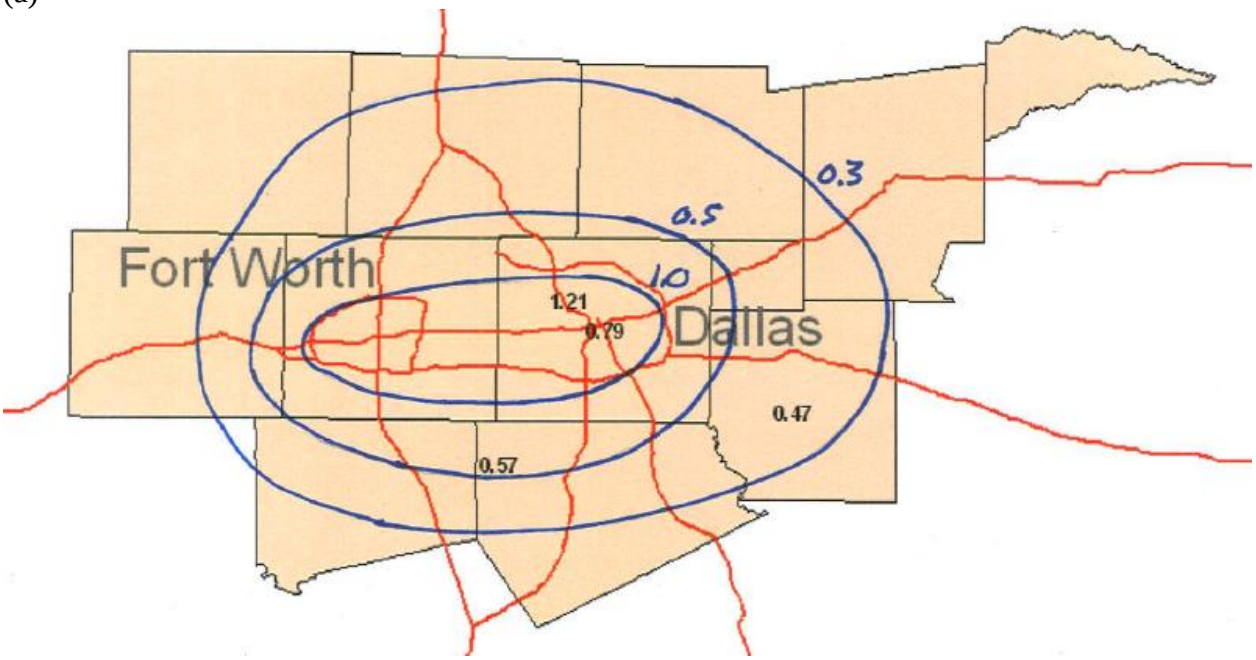


Figure 6.14. Modeled annual mean (a) and standard deviation (b) of $PM_{2.5}$-SO_4 concentration ($\mu g/m^3$) over the Dallas-Fort Worth region.

(a)



(b)

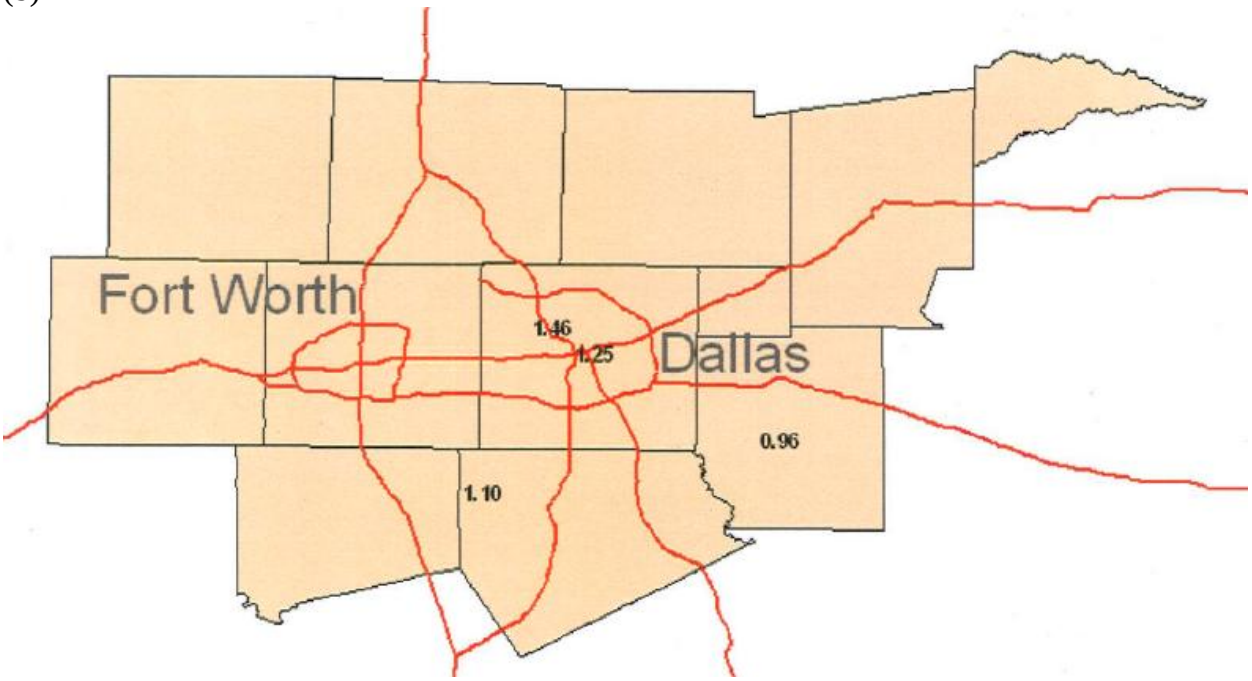
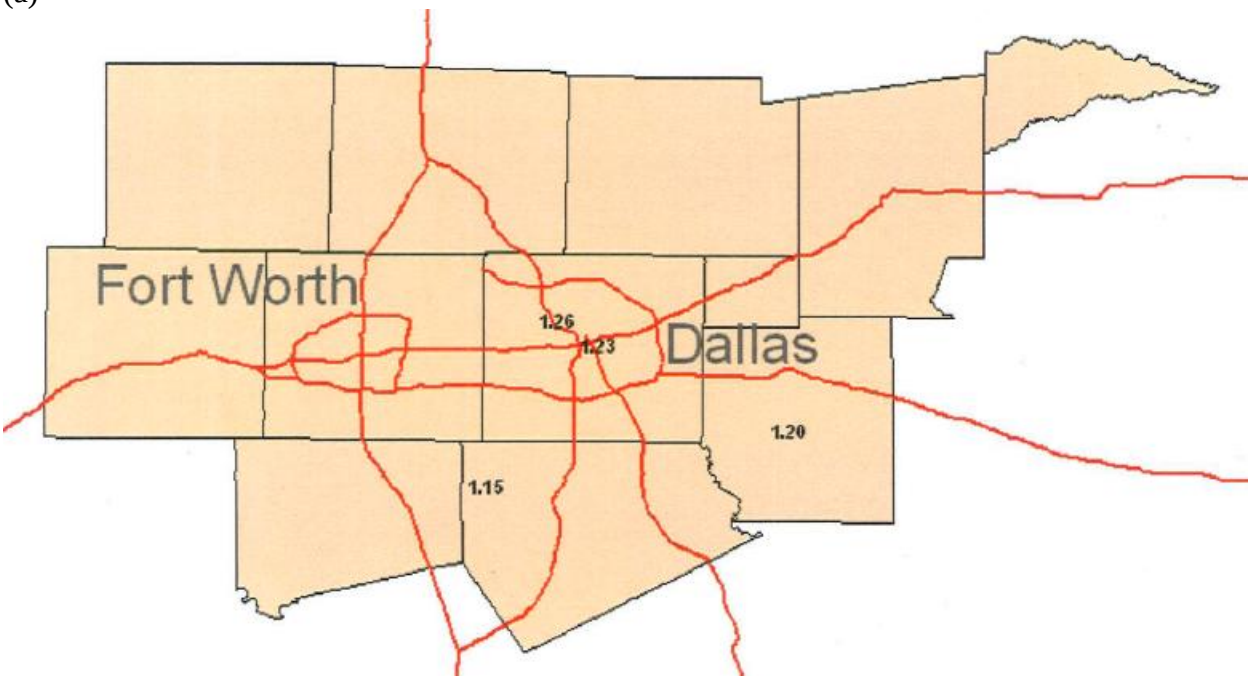


Figure 6.15. Modeled annual mean (a) and standard deviation (b) of $PM_{2.5}$-NO_3 concentration ($\mu g/m^3$) over the Dallas-Fort Worth region.

(a)



(b)

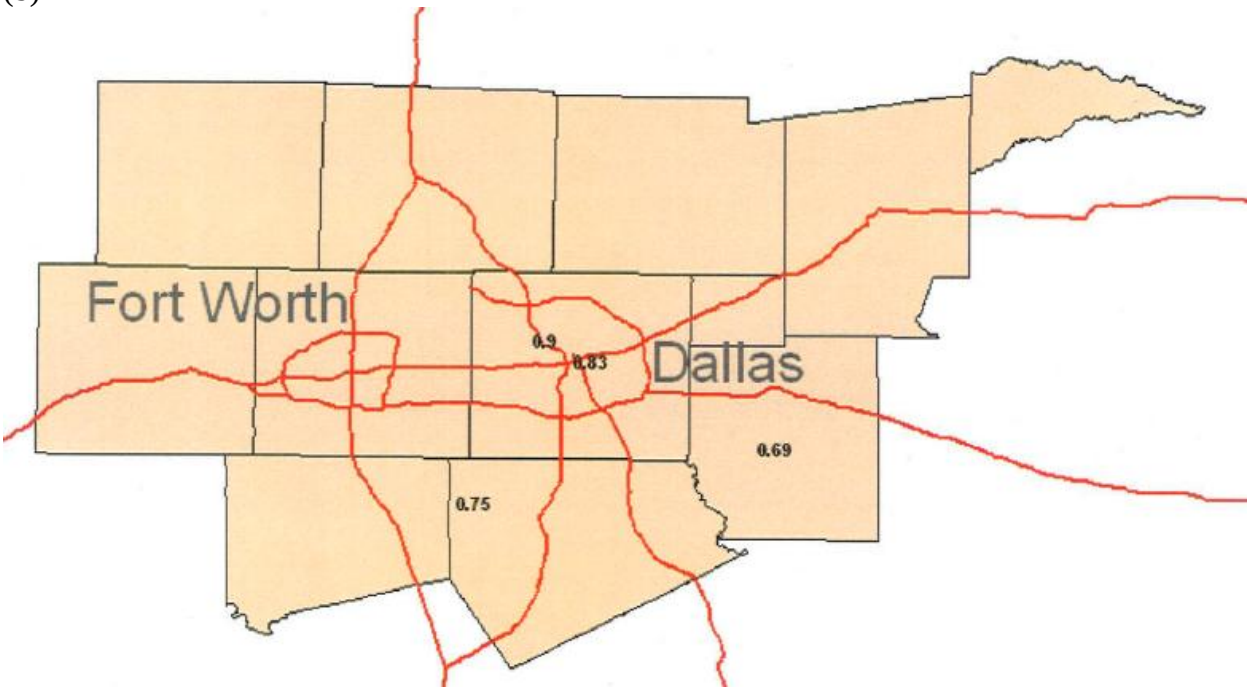
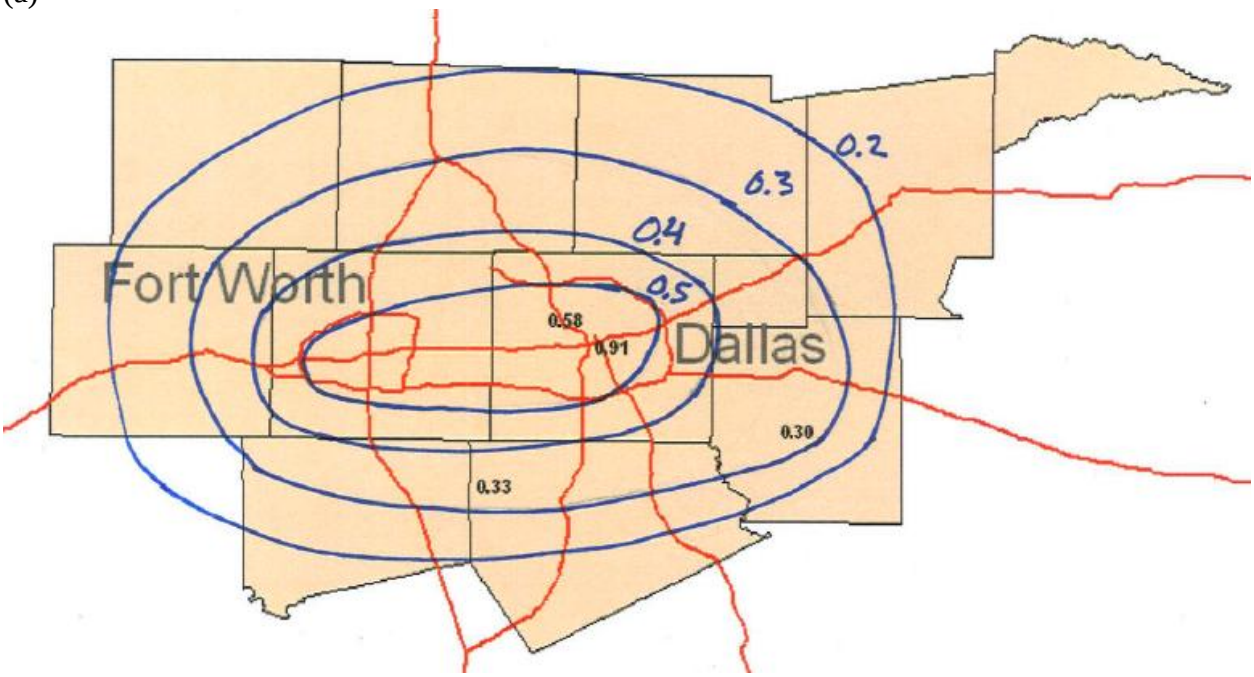


Figure 6.16. Modeled annual mean (a) and standard deviation (b) of $PM_{2.5}$-NH_4 concentration ($\mu g/m^3$) over the Dallas-Fort Worth region.

(a)



(b)

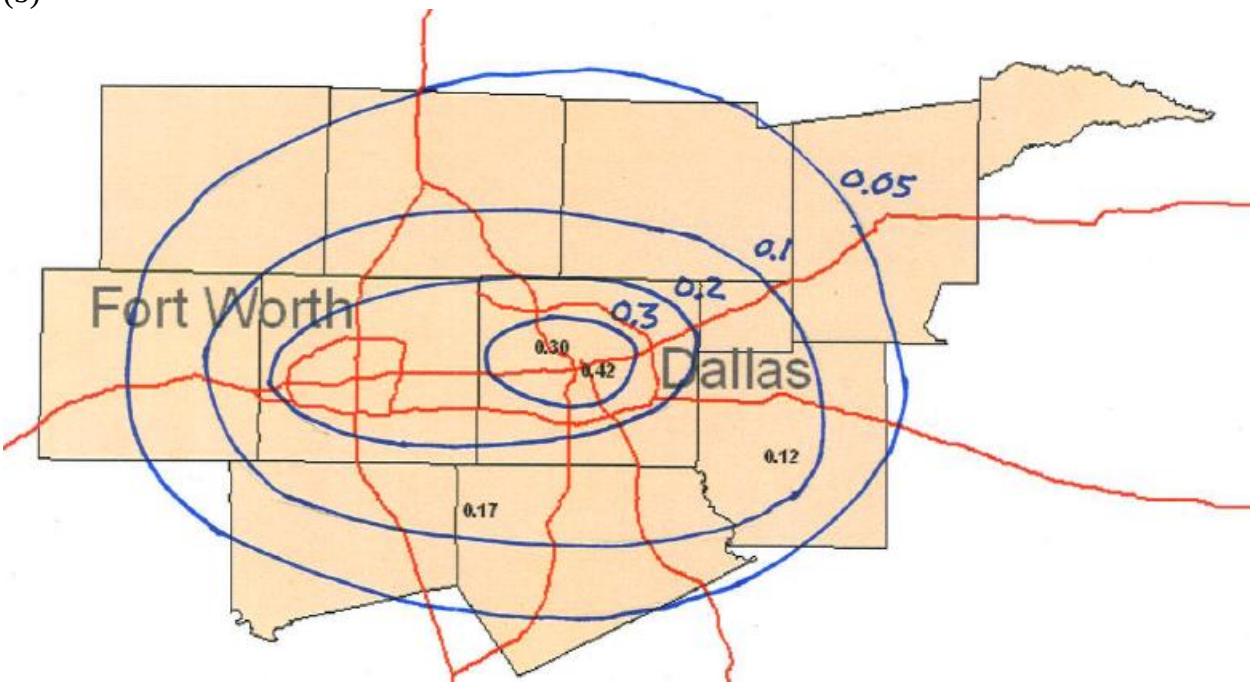
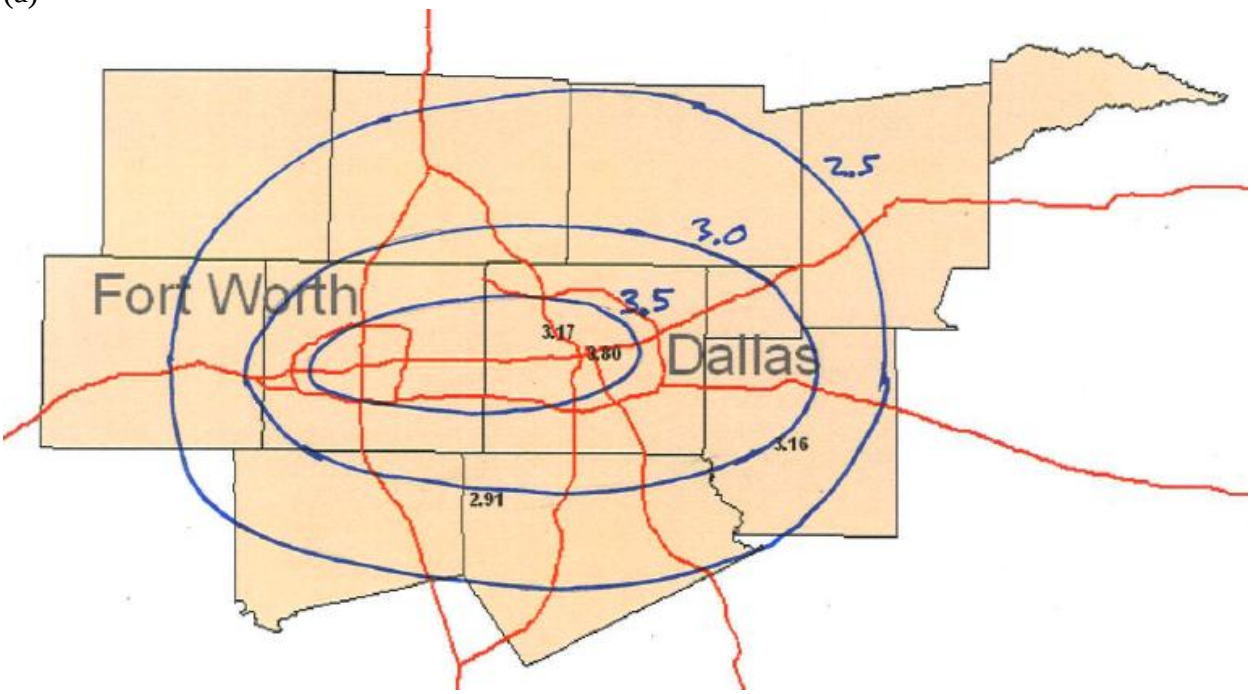


Figure 6.17. Modeled annual mean (a) and standard deviation (b) of $PM_{2.5}$-EC concentration ($\mu g/m^3$) over the Dallas-Fort Worth region.

(a)



(b)

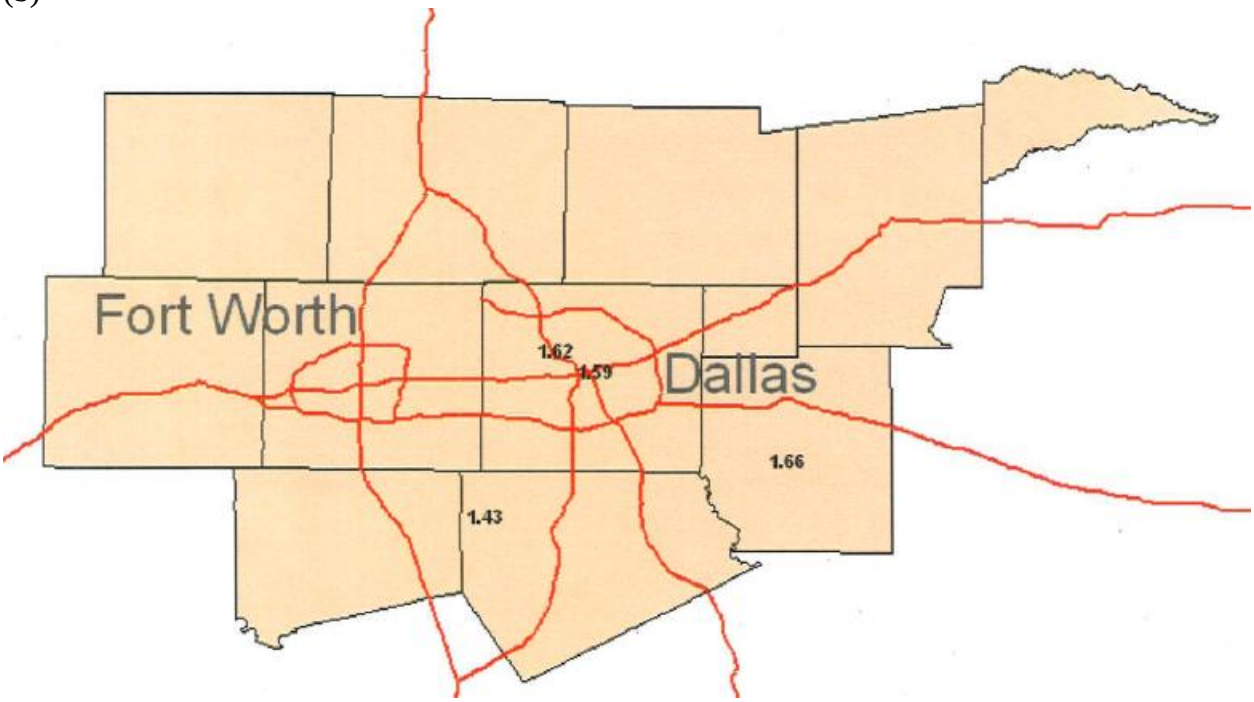


Figure 6.18. Modeled annual mean (a) and standard deviation (b) of $PM_{2.5}$-OC concentration ($\mu g/m^3$) over the Dallas-Fort Worth region.

Table 6.4. Annual mean and standard deviation of measurements at the rural Yorkville monitor in Atlanta and the estimated background concentration distribution for the Dallas-Fort Worth region.

	Atlanta Yorkville Monitor Distribution		**Dallas Estimated Background Level**	
	μ	**σ**	**μ**	**σ**
1-hr max NO_2 (ppb)	9.18	6.98	5.00	4.00
1-hr max NO_x (ppb)	10.16	9.46	5.00	5.00
8-hr max O_3 (ppb)	51.15	12.66	30.00	15.00
1-hr max SO_2 (ppb)	8.92	12.43	1.20	1.00
1-hr max CO (ppm)	0.22	0.07	0.05	0.01
24-hr PM_{10} ($\mu g/m^3$)	18.40	8.68	15.00	5.00
24-hr $PM_{2.5}$ ($\mu g/m^3$)	13.48	6.61	8.00	5.50
24-hr $PM_{2.5}$-SO_4 ($\mu g/m^3$)	4.27	3.08	3.20	2.00
24-hr $PM_{2.5}$-NO_3 ($\mu g/m^3$)	0.69	0.73	0.30	1.00
24-hr $PM_{2.5}$-NH_4 ($\mu g/m^3$)	2.06	1.31	1.20	0.60
24-hr $PM_{2.5}$-EC ($\mu g/m^3$)	0.56	0.30	0.20	0.05
24-hr $PM_{2.5}$-OC ($\mu g/m^3$)	4.87	2.25	2.50	1.50

6.3.4 Spatial Autocorrelation Assessment

To assess the spatial variability of pollutant concentrations in Dallas, correlations were calculated between the 6-year time-series at the downtown Dallas Hinton Street monitor and all other monitor time-series for each pollutant. Figures 6.19, 6.20 and 6.21 show these plots for O_3, $PM_{2.5}$ and NO_x, respectively. For secondary pollutants, correlations with the central monitor were very high even for monitors at far distances from the urban center (Figure 6.19). This trend was observed in both the Dallas and Atlanta regions. For $PM_{2.5}$ mass, correlations were also high across monitors in both cities (Figure 6.20).

For NO_x, correlations with the Hinton Street time-series decrease gradually as distance to this monitor increases. As was observed in the mean and standard deviation plots for NO_x (Figure 6.6), the Wyatt, Waterworks and Midlothian monitoring sites are anomalies. Again, it is likely that the lower correlations observed at these monitoring sites reflect their impaction by a local NO_2/NO_x source, providing further justification for

their removal from analyses which aim to capture background ambient levels, rather than variability due to local sources.

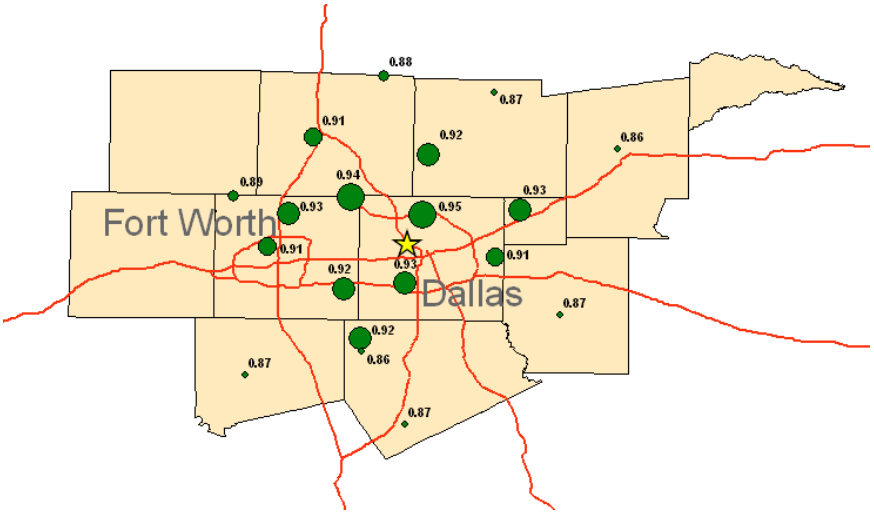


Figure 6.19. Correlation of each monitoring site time-series with the central monitor time-series for O_3 plotted at monitor locations with graduated symbols for the Dallas study area. Labels indicate Pearson correlation coefficients and yellow star indicated central monitor location.

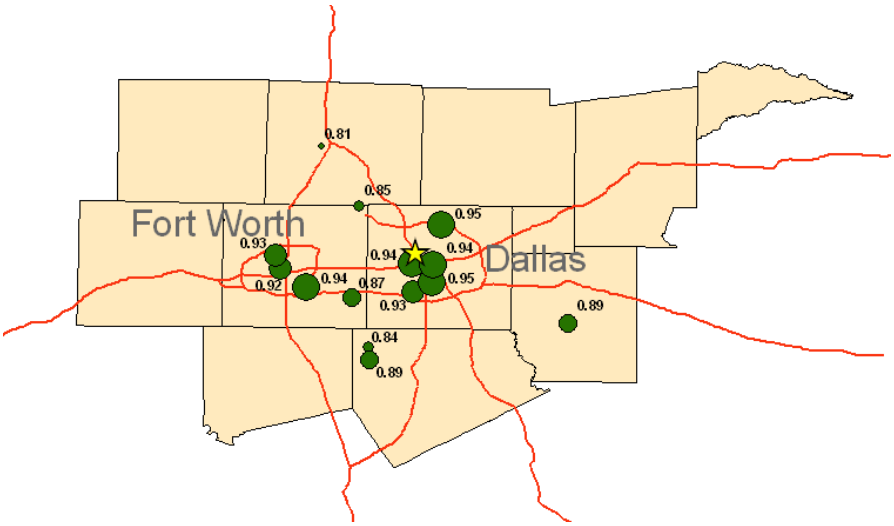


Figure 6.20. Correlation of each monitoring site time-series with the central monitor time-series for $PM_{2.5}$ mass plotted at monitor locations with graduated symbols for the Dallas study area. Labels indicate Pearson correlation coefficients and yellow star indicated central monitor location.

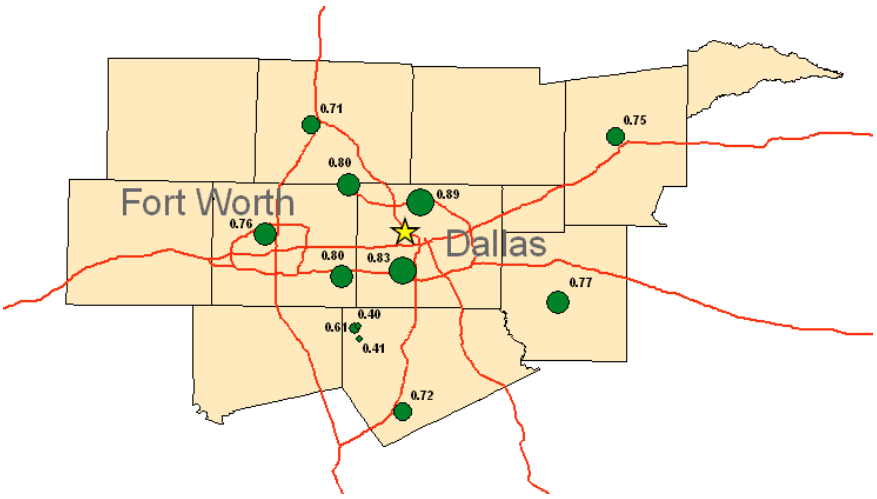
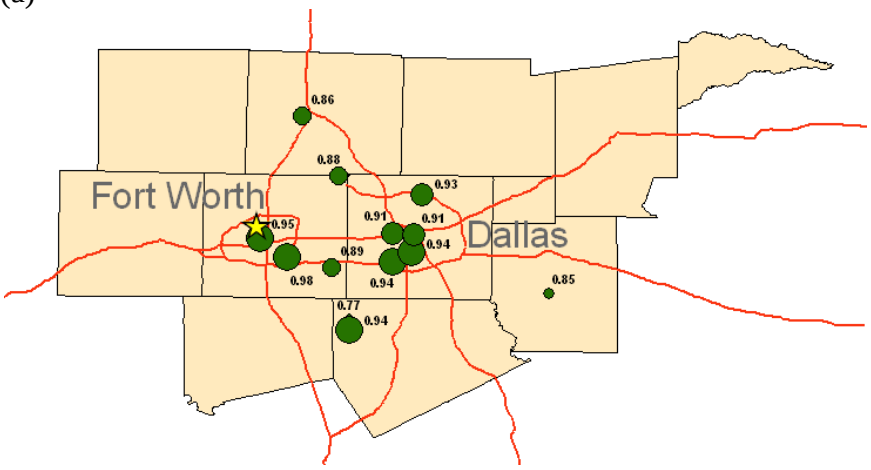


Figure 6.21. Correlation of each monitoring site time-series with the central monitor time-series for NO_x plotted at monitor locations with graduated symbols for the Dallas study area. Labels indicate Pearson correlation coefficients and yellow star indicated central monitor location.

To assess whether the correlation structure observed with the Hinton Street monitor was consistent with correlations observed for other monitoring sites, plots were constructed illustrating the correlation of all $PM_{2.5}$ mass monitor time-series with the downtown Fort Worth monitor time-series (Figure 6.22a) and with the Denton Airport South monitor, a monitor in Denton County that is removed from the two urban cores (Figure 6.22b). Correlations with the Fort Worth monitor are largely similar to those observed with the downtown Dallas site, with the highest correlations occurring with other urban monitors and lower correlations with monitors in rural locations. Accordingly, correlations are observed to decrease when time-series are compared against the Denton monitor, located more than 45 km from downtown Dallas and downtown Fort Worth (Figure 6.22b). The Denton Airport time-series has the highest correlations with other monitors removed from the city centers and the lowest correlations with monitors located close to the downtown areas of Dallas and Fort Worth; however, all observed correlations were relatively high given that $PM_{2.5}$ mass is partially derived from secondary sources and therefore, has relatively high spatial autocorrelation of measurements over the study area.

(a)



(b)

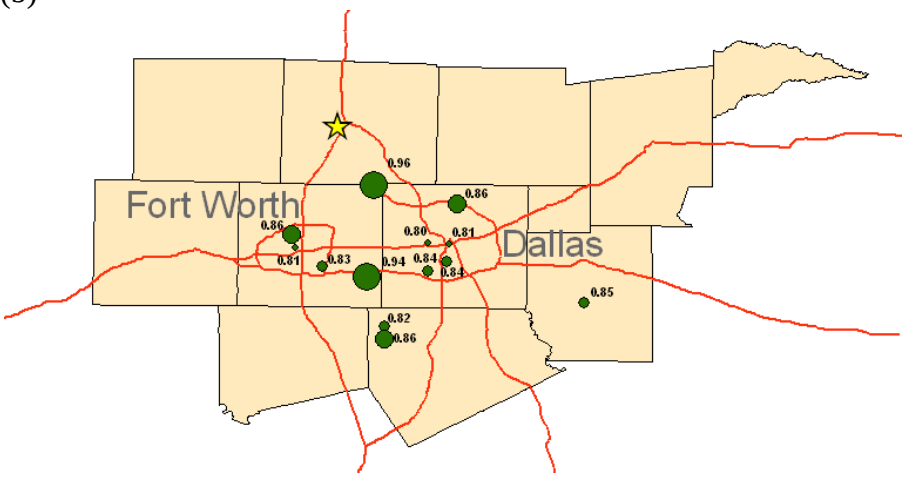


Figure 6.22. Correlation of each monitoring site time-series with the downtown Fort Worth time-series (a) and with the Denton Airport South monitor (b) for $PM_{2.5}$ mass plotted at monitor locations with graduated symbols. Labels indicate Pearson correlation coefficients and yellow star indicated the location of the Fort Worth monitor (a) or the Denton Airport South monitor (b). Symbol sizing is consistent across plots.

Semivariograms were constructed for the 12 ambient air pollutants (Figure 6.23). All monitor pairs with time-series of greater than 365 coinciding measurements during the study period of 2003-2008 were included in the semivariogram analysis. For this assessment, the spatial variance of pollutants was treated as isotropic. Exponential theoretical semivariogram curves were regressed to the raw semivariogram values, constraining the sill to 1 and the nugget to a value representative of pollutant-specific instrument precision error (Goldman et al 2010).

The semivariogram analysis aims to characterize the variability due to the regulatory ambient, that is, the variability that is representative of ambient concentrations, not variability associated with concentrations near sources. As such, measurement data from monitors believed to be impacted by local sources were excluded from this analysis. For $PM_{2.5}$ mass, the White Settlement monitor was excluded due to the monitor's location 100 m from Interstate 30 and the fact that it exhibited uncharacteristically low correlations with other $PM_{2.5}$ mass monitoring sites in the region, both suggesting that measurements from this monitor are likely to be traffic impacted. For NO_2 and NO_x, the Wyatt, Waterworks and Midlothian monitors were not included in the semivariogram analysis as their elevated concentrations and poor inter-correlation despite close proximity suggest that these monitors are heavily influenced by a local industrial source.

For SO_2, a semivariogram could not be effectively estimated due to the highly variable nature of the pollutant and limited number of monitoring sites available. Because SO_2 is emitted largely from point sources, its distribution over space is based largely on plume path and wind direction. As a result, any spatial model derived from monitoring data is very sensitive to monitor placement relative to these point sources. In this analysis, six monitoring site locations provide data for SO_2 for this time-period. Inter-monitor correlations suggested the location of an SO_2 point source between monitoring site locations as several correlations were less than zero. Table 6.5 lists the inter-monitor correlations for SO_2 measurements.

Table 6.5. Distance between sites and inter-monitor correlations for SO_2 measurements

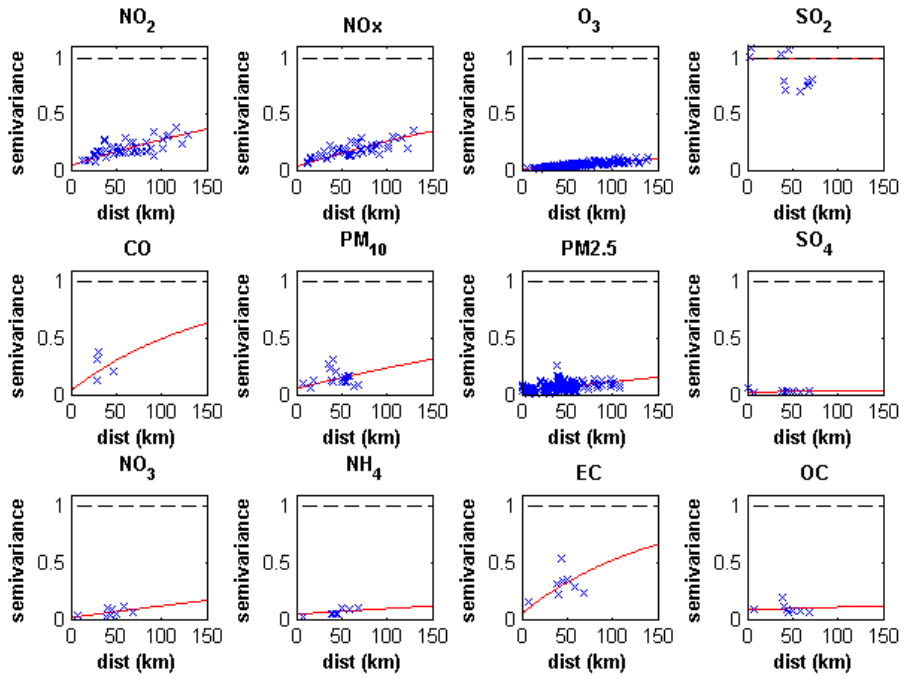
Monitor Pair		distance (km)	R
Hinton	Italy	71.29	0.11
Hinton	Wyatt	42.04	0.17
Hinton	Midlothian OFW	40.58	0.11
Hinton	Waterworks	45.20	-0.04
Hinton	Kaufman	58.14	0.18
Italy	Midlothian OFW	36.84	-0.02
Italy	Kaufman	67.35	0.11
Wyatt	Midlothian OFW	1.74	0.00
Wyatt	Waterworks	4.38	-0.04
Wyatt	Kaufman	68.64	0.12
Midlothian OFW	Waterworks	5.02	-0.09
Midlothian OFW	Kaufman	67.06	0.14
Waterworks	Kaufman	67.76	-0.14

Based on the Dallas semivariograms (Figure 6.23a), traffic-related pollutants (NO_2, NO_x, CO, $PM_{2.5}$-EC) show a greater degree of spatial variability, with larger variation between monitors even at relatively short distances. In contrast, O_3, $PM_{2.5}$ mass and $PM_{2.5}$ components SO_4, NO_3, and NH_4 show low levels of variability over space even at greater distances between monitors. PM_{10} and $PM_{2.5}$-OC exhibit intermediate levels of spatial variability.

Dallas semivariogram results are compared to those from a previous study in Atlanta (Goldman et al. 2010) (Figure 6.23b). Because the nugget and sill values were the same in the semivariance curves constructed for Dallas and Atlanta, the range values provide a direct measure to compare the two cities (Table 6.6). Longer range values indicate greater spatial homogeneity, while smaller range values indicate higher spatial variability. Population-weighted semivariance values, γ_{pw}, were computed for both cities using census tract population data and provide an intercity comparison of spatial variance for each pollutant (Table 6.6). Semivariograms for Dallas and Atlanta are similar in some regards with both indicating higher spatial variability for pollutants that are largely emitted to the atmosphere from sources directly (NO_2, NO_x, CO, $PM_{2.5}$-EC) and greater

spatial homogeneity for pollutants formed largely through secondary formation processes (O_3 and $PM_{2.5}$ components SO_4, NO_3, and NH_4); however, there are several differences of note. All population-weighted semivariances computed for pollutants in Dallas are lower than those in Atlanta, indicating that greater spatial homogeneity is observed in Dallas for all pollutants. In particular, NO_2 and NO_x demonstrate less variability in Dallas than Atlanta with longer range values regressed and considerably lower population-weighted averages. It should be noted, however, that this semivariogram analysis is highly sensitive to the number and placement of monitoring sites; therefore, pollutants with fewer available monitors (e.g. CO in Dallas) have a higher degree of uncertainty in their variance models. Scatter observed in the raw semivariances can likely be attributed to anisotropic variance effects as well as local emission and meteorological influences.

(a)



(b)

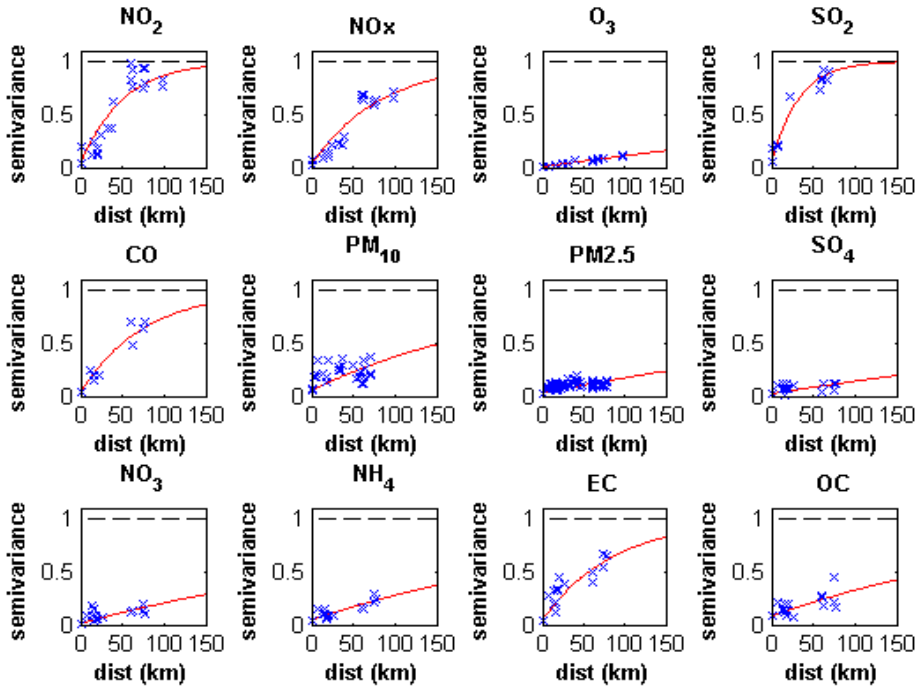


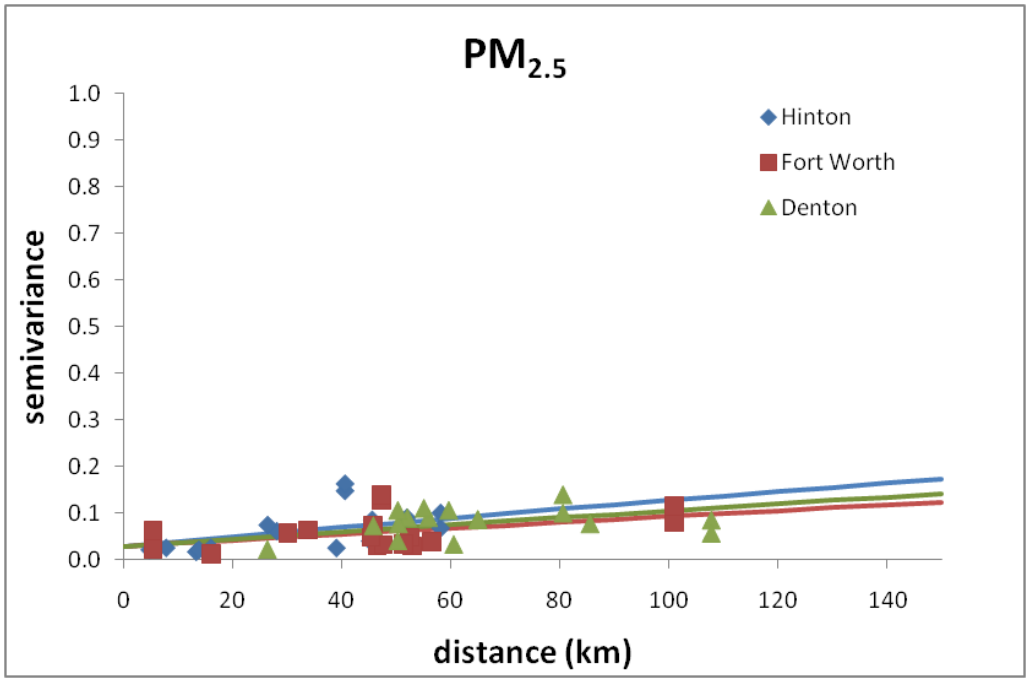
Figure 6.23. Semivariograms for (a) Dallas and (b) Atlanta for normalized pollutant concentrations. Raw semivariance values are plotted (blue) and exponential curves (red) are fitted to the data.

Table 6.6. Range (km) from semivariogram function and population-weighted semivariance, γ_{pw} for Dallas and Atlanta semivariogram models.

	Dallas		Atlanta	
	range	γ_{pw}	range	γ_{pw}
1-hr max NO_2	120.52	0.127	17.05	0.461
1-hr max NO_x	125.81	0.115	28.25	0.326
8-hr max O_3	501.97	0.030	303.97	0.043
1-hr max SO_2	-	-	10.89	0.596
1-hr max CO	51.95	0.223	25.69	0.352
24-hr PM_{10}	155.33	0.120	81.26	0.170
24-hr $PM_{2.5}$	367.02	0.059	208.15	0.079
24-hr $PM_{2.5}$-SO_4	2852.65	0.027	258.36	0.063
24-hr $PM_{2.5}$-NO_3	300.54	0.050	153.69	0.080
24-hr $PM_{2.5}$-NH_4	626.53	0.061	119.04	0.126
24-hr $PM_{2.5}$-EC	48.54	0.242	27.92	0.340
24-hr $PM_{2.5}$-OC	1252.79	0.091	108.38	0.169

To assess the sensitivity of the semivariogram analysis to spatial location, additional semivariograms were constructed for NO_x and $PM_{2.5}$, in which semivariance curves were regressed based on monitor-specific correlations. Figure 6.24 illustrates the raw semivariances and regressed curves based on correlations for three monitoring sites: the Hinton Street monitor in downtown Dallas, the downtown Fort Worth monitoring site and the Denton Airport South monitoring site. Only small differences between the curves derived from each of these sites are observed, suggesting that the semivariance model used here is not sensitive to spatial location. This demonstrates that while the anisotropic structure of population and pollution variability observed in Dallas may have important implications for spatial modeling of concentrations in two dimensions, the one dimensional semivariogram analysis performed here is relatively robust and thus, application of an isotropic variance model is reasonable for estimation of spatial autocorrelation for ambient air pollutants in the Dallas-Fort Worth region.

(a)



(b)

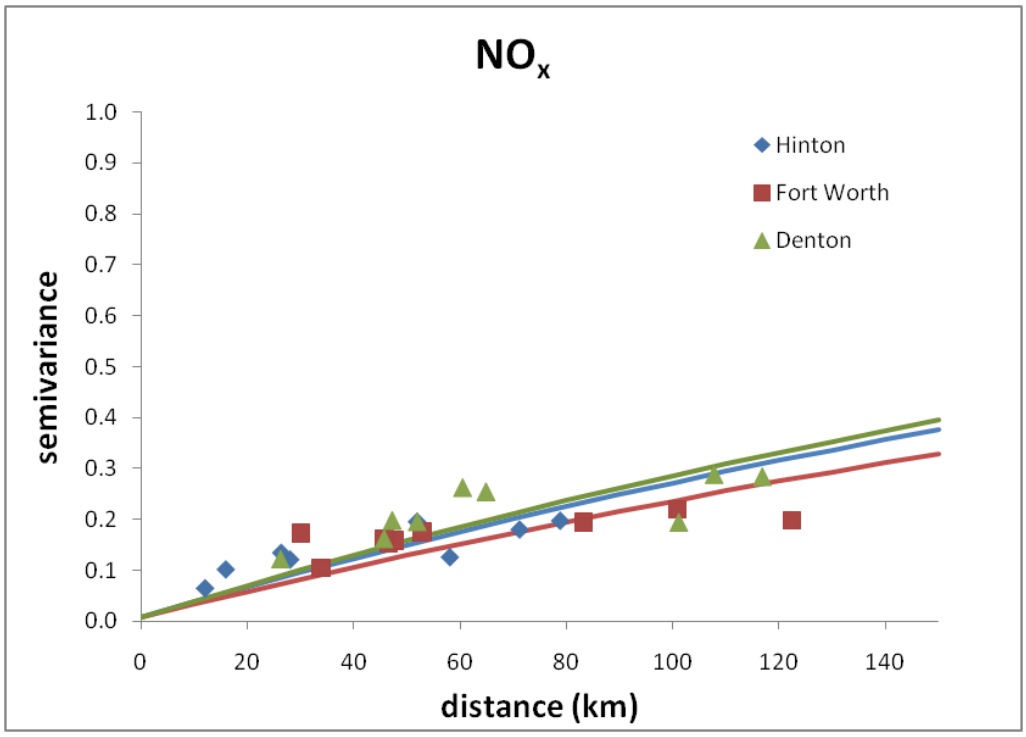


Figure 6.24. Semivariance models for (a) $PM_{2.5}$ mass and (b) NO_x derived from correlations of all monitoring sites with three different monitor time-series: Hinton, Fort Worth and Denton.

To assess the sensitivity of semivariogram curves to the raw semivariance values used to construct the semivariance models, curves were regressed for a 5% increase and 5% decrease in the semivariance values between each monitor pair for each pollutant (Figure 6.25). Results indicate that the semivariogram curves are relatively robust to changes in the variance between monitoring sites.

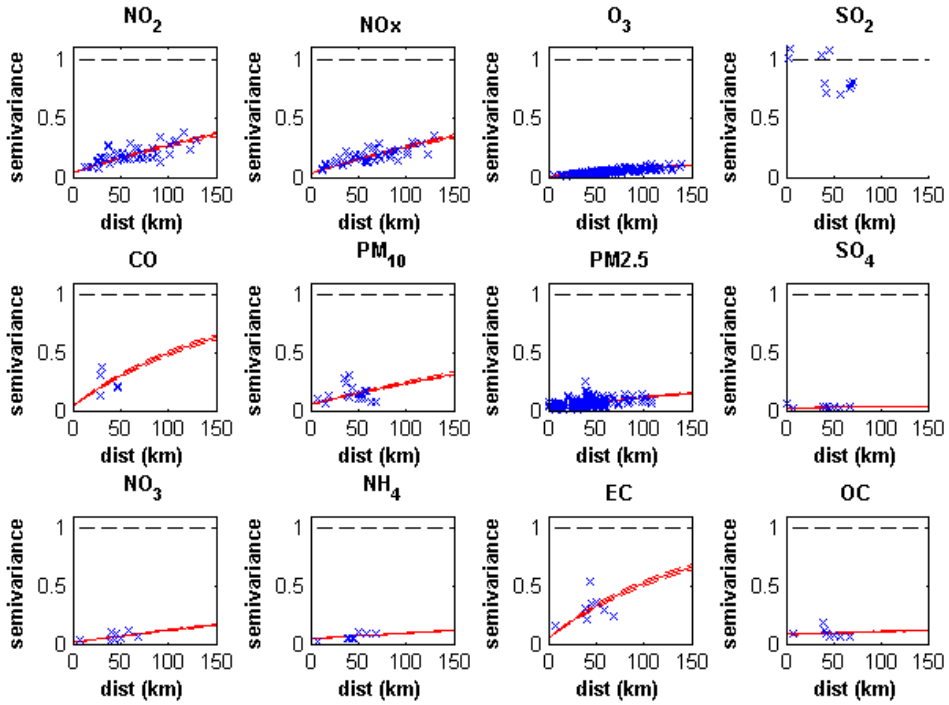


Figure 6.25. Regressed semivariogram curves (solid red) ± 5% of semivariance values for monitor pairs (dashed red) used to regress the semivariance models.

6.4 Discussion

This analysis has provided some insight into the spatio-temporal distribution structure of 12 ambient air pollutants in the metropolitan region surrounding Dallas, Texas. Assessment of the mean and standard deviation of monitoring site time-series over space for the Dallas-Fort Worth region indicates that an anisotropic model of pollutant concentrations may be most appropriate for spatial modeling of pollutant distributions for this region. In Atlanta, population and source distributions largely decrease uniformly in

all directions from downtown; thus, isotropy of pollutant concentrations may be a more reasonable assumption in this city. By contrast, Dallas is not an isolated urban center to the degree that Atlanta is. The presence of the twin city Fort Worth and population centers in between the two urban hubs, such as Arlington and Irving, create an elongated population distribution pattern in the east-west direction; thus, air pollutant concentrations, particularly those that increase proportionally with population such as traffic-related species, are not likely to be isotropic with respect to downtown Dallas. Consequently, a more sophisticated model of pollutant concentration is advised in order to fully characterize the distribution of ambient air pollutants in the Dallas-Fort Worth region.

In the context of health studies of Dallas, it will be important to consider the observed elliptical pattern of population and pollutant concentration in this region. Any population-weighting of pollutant concentrations will likely need to be done on a census tract (or other discrete geographic unit) basis rather than through use of a continuous function with respect to distance from the urban center, as has been done for Atlanta in past work (Goldman et al. 2010). This will be necessary in order to better capture the asymmetric population distribution. The unique population spatial pattern in the Dallas-Fort Worth area may also have important implications for the personal exposure assessment, as typical assumptions about spatial distribution of pollutants as well as individual time-activity patterns may not be reasonable for this region. Any health study design that considers spatially resolved air pollutant or population data should take thorough consideration of the observed anisotropic distribution of pollutant concentration in the Dallas-Fort Worth region.

In terms of correlation between monitors over space, it was found that Dallas could be modeled isotropically. Semivariograms were produced for all 12 pollutants and a sensitivity analysis of NO_x and $PM_{2.5}$ semivariograms suggest that the variance structure modeled in the semivariogram plots is relatively independent of location in

space. This suggests that an isotropic variance model is a reasonable approximation for the characterization of spatial autocorrelation in the Dallas-Fort Worth region.

The influence of point sources in the analysis of spatial variability is also a central concern. Because point sources emit pollutants directly to the atmosphere from a single geographic location, the spatial distribution of ambient pollutants originating from these sources is not well characterized by a small number of fixed outdoor monitors, since the ability of these monitors to capture these sources is dependent on meteorological conditions and expected to be highly variable. The influence of point sources on the spatial distribution of ambient concentrations may be significant for some species, as was demonstrated here for SO_2. If point sources are found to be introducing high variability in ambient concentrations, the ability to find significant health associations may be severely limited for these pollutants (Goldman et al. 2010). Health assessment for pollutants emitted from point sources or those otherwise exhibiting high spatial variability should consider the impact of this variability on results and its potential to limit findings of significant associations.

Further, previous studies have suggested that the impact of spatial variability error on epidemiologic studies is modified depending upon the type of error present in the data (Goldman et al. 2011; Zeger et al. 2000). Determination of error type in a given dataset is challenging and may vary considerably for different pollutants and metrics used. Additional assessment of spatial variability of ambient concentrations in the Dallas metropolitan region will be necessary to assess error type of pollutant distributions and its impact on time-series health studies.

The spatio-temporal distributions of 12 ambient air pollutants around the Dallas metropolitan area were found to be slightly more homogeneous over space than pollutants in Atlanta; therefore, the impact of air pollutant measurement error due to spatial variability on a time-series health study in Dallas is estimated to be less than the impact assessed for Atlanta. General comparative results from Dallas across pollutants

are expected to agree with Atlanta findings, however, with greater risk estimate attenuations observed for more spatially heterogeneous pollutants and smaller attenuations for more spatially homogeneous pollutants (Goldman et al. 2010; Goldman et al. 2011).

6.5 Conclusions

This assessment has provided an initial analysis on the spatio-temporal distribution of 12 ambient air pollutants in the area surrounding Dallas, Texas for the time period of 2003-2008. Monitor coverage varies by pollutant with a minimum of two monitors with measurement for the duration of the study period. Daily metrics are available for gases (NO_2, NO_x, O_3, SO_2, and CO) and $PM_{2.5}$ for the entire 2003-2008 time period. Daily metrics of $PM_{2.5}$ components (SO_4, NO_3, NH_4, EC, and OC) are available for 2006-2007. PM_{10} mass is available at a frequency of six days for the duration of the study period. The spatial distribution and correlation of pollutants was characterized. The dual urban centers of Dallas and Fort Worth create a more complex spatial structure of concentration distributions than is observed in Atlanta; thus, anisotropic models for mean and standard deviation of concentrations would better approximate pollutant distributions for the region than an isotropic model. Semivariograms were constructed and a range of variability was observed across pollutants, with primary pollutants showing greater degrees of variability over space and secondary pollutants showing greater correlation between monitors over space. A sensitivity analysis demonstrated that semivariogram models varied little when assessed at different locations in space, suggesting that an isotropic variance model is reasonable for the characterization of spatial autocorrelation for this region. Sulfur dioxide is particularly difficult to characterize in Dallas, given monitor placement and point sources in the region. Overall, measurement error due to the spatial variability of pollutants observed in Dallas is estimated to result in less attenuation

of health associations than that assessed in a time-series study of Atlanta air pollution and acute health effects.

CHAPTER 7

CONCLUSIONS AND FUTURE DIRECTIONS

Ambient air pollution in urban areas is regulated largely through use of fixed outdoor monitoring sites intended to capture ambient levels of pollutants. Measurements from these monitors are used as surrogates for population exposure to ambient concentrations in health studies. Time-series epidemiologic studies utilize measurements from these monitors to examine correlations between air pollution levels and changes in daily health outcome variables. The strength of association found in these studies is shown to be influenced by the spatial variability and monitor network of the pollutant of interest. This dissertation explores the impact of this air pollution measurement error on time-series epidemiologic studies. In Chapter 3, the amount of air pollution measurement error due to instrument precision and spatial variability was characterized and its impact on a time-series study of cardiovascular disease emergency department visits in Atlanta was assessed by adding error to observations. In Chapter 4, understanding of this measurement error was enhanced further through assessment of the modification of the impact of this error by error type. Results were illustrated for a range of error types from Berkson error to classical error defined on a log scale, providing a thorough assessment of potential impacts on health associations expected from error of these types. While Chapters 3 and 4 characterized measurement error amount due to instrument imprecision and spatial variability of ambient pollutants and quantified the range of expected biases in risk estimates in a time-series study due to this measurement error being of different types, these chapters did not specify the type of error present in actual ambient concentrations, nor did they characterize the amount of measurement error resulting from spatial heterogeneity of ambient concentrations. Thus, to build on the methodologies developed in Chapters 3 and 4, these issues are addressed in the following chapter. In

Chapter 5, ambient concentrations were simulated over space and time and the concentration fields were generated to assess the amount and type of error due to spatial variability in true ambient concentrations, as well as to quantify the impact of use of different monitor-based air pollutant exposure metrics in a simulated time-series study. This chapter allowed for assessment of the amount and type of error actually present in ambient concentrations because it provided an estimate of the true variability in ambient concentrations over space and time. With thorough methodologies for the characterization of measurement error and its impact on health studies established for Atlanta in Chapters 3, 4 and 5, the research methods developed could be applied to ambient air pollution measurements in a different metropolitan area. In Chapter 6, the measurement error analyses developed in Atlanta were applied to a time-series study in Dallas, Texas. Measurement error due to spatial variability in Dallas was assessed and the expected impact of this error on a time-series epidemiologic study was discussed.

Overall, this dissertation has demonstrated that measurement error due to instrument imprecision and error resulting from spatial variability and spatial heterogeneity of ambient concentration can be effectively characterized using measurement data from fixed outdoor monitoring sites and the impact of these measurement errors on health risk estimates in time-series studies can be quantitatively assessed. Methodologies were developed to specify the amount and type of error present in measurements and geostatistical simulation can aid in informing on the amount and type of error actually present in true ambient concentrations. Methodologies developed are applicable to ambient monitor networks in other cities. Given below are conclusions from individual chapters in this dissertation.

Ambient Air Pollutant Measurement Error: Characterization and Impacts in a Time-Series Epidemiologic Study in Atlanta

Ambient air pollution measurement error from instrument imprecision and spatial variability was characterized for Atlanta monitoring sites and the impact of these errors on a time-series epidemiologic study of cardiovascular disease emergency department visits was assessed. Instrument error was found to have little impact on the health effect estimate, while error due to spatial variability resulted in greater attenuation of risk estimates, particularly for primary pollutants which have greater amounts of spatial heterogeneity. These observations suggest that it is more likely to observe positive associations in secondary pollutants than primary pollutants since secondary pollutants are measured with less error due to less spatial variability.

Impact of Exposure Measurement Error in Air Pollution Epidemiology: Effect of Error Type in Time-Series Studies

The influence of error type on the impact of air pollution measurement error on time-series studies was examined through simulation of air pollutant time-series over a range of error types from classical on a log scale to Berkson on a log scale. Error type was observed to have a profound effect on the bias due to measurement error in the time-series study. While the introduction of measurement error as purely classical on a log scale resulted in significant biases toward the null hypothesis, the addition of measurement error as purely Berkson on a log scale resulted in bias away from the null, with the introduction of error leading to a stronger association observed, albeit less significant. The type of error present in a measurement dataset greatly impacts the associations observed.

Quantification and Impacts of Spatial Variability Error in Time-series Studies of Ambient Air Pollution: A Simulation Study

A spatio-temporal model was developed to simulate ambient air pollutant concentrations over the Atlanta metropolitan area for a 6-year period. The simulated concentration fields were used to quantify the amount of error resulting from spatial variability in ambient concentrations, as well as from instrument error associated with measurement imprecision and the number and placement of monitors. To address the latter, simulations of monitoring site data were generated to explore the amount and type of additional error introduced through use of monitor-based exposure metrics. Most of the error assessed was found to be due to the spatial variability of ambient concentrations, rather than error introduced through use of monitoring site measurements. Error due to spatial variability alone was found to be Berkson error and thus, is expected to yield no bias towards the null hypothesis, whereas error with instrument error included is predicted to result in some attenuation of risk estimates. The simulated ambient concentration fields are being used to generate simulated health outcome data in order to assess the impact of spatial variability on risk estimates and significance levels in an epidemiologic model.

Spatio-temporal Variability of Ambient Air Pollutants in Dallas, Texas: Implications for Time-Series Epidemiologic Studies

Air pollution measurement error due to spatial variability for the Dallas metropolitan area was assessed using methods developed from an ongoing study of air pollution and health impacts in Atlanta. Analysis of 12 ambient air pollutants suggested that the methodologies developed for Atlanta are translatable to Dallas. Results indicate that most pollutants, particularly NO_2, and NO_x are more spatially homogeneous in Dallas than in Atlanta. As such, attenuations in health risk estimates due to spatial variability of ambient concentrations is anticipated to be less than those estimated in

Atlanta. In both Atlanta and Dallas, SO_2 is more difficult to characterize due to the location of point sources and the limited number of monitors.

Future Directions

Semivariogram Analysis

In this work, the quantity of measurement error due to spatial variability was assessed largely through semivariogram analysis. While this tool provides useful information on the variability of ambient concentrations over space, it has limitations in its scope. First, the semivariogram analysis used here assumes isotropic behavior of air pollutants; that is, the spatial autocorrelative structure is assumed to be same in all directions. Although this assumption is reasonable and appropriate in some locations and for some pollutants, such as in Atlanta and Dallas for the 12 pollutants studied, more complex spatial correlation structures in ambient concentrations likely exist in other locations that are not captured by this simplification. Consequently, future work could involve development of a more sophisticated model of spatial autocorrelation, in which anisotropic effects are considered. While such an analysis would be limited by the number and location of monitoring sites available, more informed spatial models could be developed for pollutants with greater numbers of monitoring sites. Further, additional tools such as land use regression modeling, plume modeling for pollutants emitted from sources primarily and chemical mass transport modeling could be utilized to provide additional information on the directionality and location-specific differences in semivariance. An assessment of measurement error due to spatial variability that accounts for anisotropic effects could enhance the work that has been done here.

Geostatistical Simulation of Air Pollutant Concentrations

In Chapter 5, simulated air pollutant concentration fields were generated to represent variability in ambient concentrations over space and time using the Stanford

Geostatistical Modeling Software (SGeMS). Monitoring site time-series were simulated that included instrument error and simulations were used to assess the error due to spatial variability of ambient concentrations and error due to the use of monitor-based metrics in time-series studies. The concentration time-series simulated in this study could be used to further explore the influence of the number and location of monitoring sites on measurement error and health impact assessment. The sensitivity of results to the number and location of monitors could be assessed by adding and removing simulated monitor time-series and by simulating theoretical monitor time-series at different locations in space. The simulated monitor time-series could then be compared to the true ambient concentration fields produced in SGeMS and the measurement error could be characterized for different scenarios of number and locations of monitors. Such an analysis would be of great value for air pollution regulatory agencies and health scientists looking to optimize the number and location of monitors to best capture ambient concentration variability or population exposure.

In Chapter 5, only spatial and temporal variability in the background, or "regulatory" ambient concentrations were simulated in the true ambient concentration fields; that is, only concentration variability that would be expected from ambient monitors not impacted by local sources directly was reproduced. Any microscale variability of concentrations over time or space, such as that associated with sub-hour meteorological events or concentration variability in close proximity to sources, was not modeled. Further work could include simulation of concentrations representative of these additional levels of concentration variability and the associated amount and type of error could be characterized. Such assessment would be informative for understanding total ambient exposure of populations or looking at health effects associated with peak concentrations on short time scales.

Another possibility is the simulation of other kinds of air pollution datasets, such as indoor air quality or individual exposures. Simulations could be produced, for

example, that would include variability in exposure of individuals as they move through time and space. Such modeling could account for personal exposures in microenvironments such as in-vehicle, cooking and occupational exposures. The type of error associated with variability in individual exposures also could be assessed. Such analyses would inform on the influence of the amount and type of error associated with these kinds of datasets and shed additional light on the best choice of exposure metric for use in different epidemiologic study designs.

While many studies have discussed the concept of error type and its qualitative impacts on epidemiologic studies, less work has focused on the determination of error type for specific datasets and the quantification of impacts, due to the fact that true ambient concentrations are unknown. Chapter 4 demonstrated impacts of type on a time-series study of measurement error evaluated over a range of error types. In Chapter 5, simulated time-series of ambient concentrations were generated to represent variability in ambient concentrations over space and time. The fact that true concentration fields were estimated allowed for the determination of the error type associated with spatial variability of ambient concentrations, as well as various pollutant exposure metrics, since error type assessment relies on some approximation of true concentrations. Additional analyses could be done to supplement this work by assessing error type for different kinds of air pollutant datasets, such as those discussed above, including microscale variability in ambient concentrations and personal exposures.

Assessment of Measurement Error in Different Cities

Chapter 6 provided an initial assessment of the measurement error associated with ambient air pollutant concentrations from monitoring site networks in the area surrounding Dallas, Texas for a time-period of 2003-2008. The analysis will inform an epidemiologic study that is currently being conducted for this region; however, additional assessment of exposure measurement error for this study will be of value. First, a full

investigation of emission sources impacting the region will be necessary. Beyond the anthropogenic emissions typical of urban areas such as those from mobile sources, Dallas has several industrial and agriculture operations in the region that may be significant sources of some species of epidemiologic interest in this study. Further research on the frequency, quantity and other emission characteristics of these sources would provide additional insight into the level of spatial variability for impacted pollutants.

Additionally, work could be done to simulate Dallas concentrations over time and space using SGeMS, as was done for Atlanta in Chapter 5. These simulations could apply the anisotropic spatial models of annual mean and standard deviation of concentrations proposed for the Dallas-Fort Worth region in Chapter 6. Simulated ambient concentration fields of this region that accounted for the elliptical population and pollution pattern observed would be informative for assessment of air pollutant measurement error for the region, as well as for characterization of population exposure to aid in the current health study.

Ambient air pollution measurement error could also be assessed in other cities. Dallas was observed to have a unique population and air pollutant distribution due to the dual urban centers of Dallas and Fort Worth that present a modeling challenge for air pollution and health studies. Other cities also have different geographic or demographic features that could influence population and air pollutant concentrations. For example, coastal cities present an interesting challenge in that population and air pollutant concentrations are likely not to have symmetric distributions with respect to the urban center. Orographic features also influence population distribution for cities in mountainous regions. These features could have significant impacts on population exposure to ambient air pollution and have important implications in epidemiologic studies. More complex models of pollutant concentration over space that account for these anisotropic effects could be developed in order to characterize air pollutant measurement error in cities with these features and inform potential health studies. The

type of error present in exposure metrics for other metropolitan area could be assessed as well, as it is feasible that findings may differ from city to city, given observational differences in pollutant distributions and geographic characteristics discussed above.

APPENDIX A

SUPPLEMENTAL MATERIAL FOR INSTRUMENT AND SPATIAL ERROR ASSESSMENT

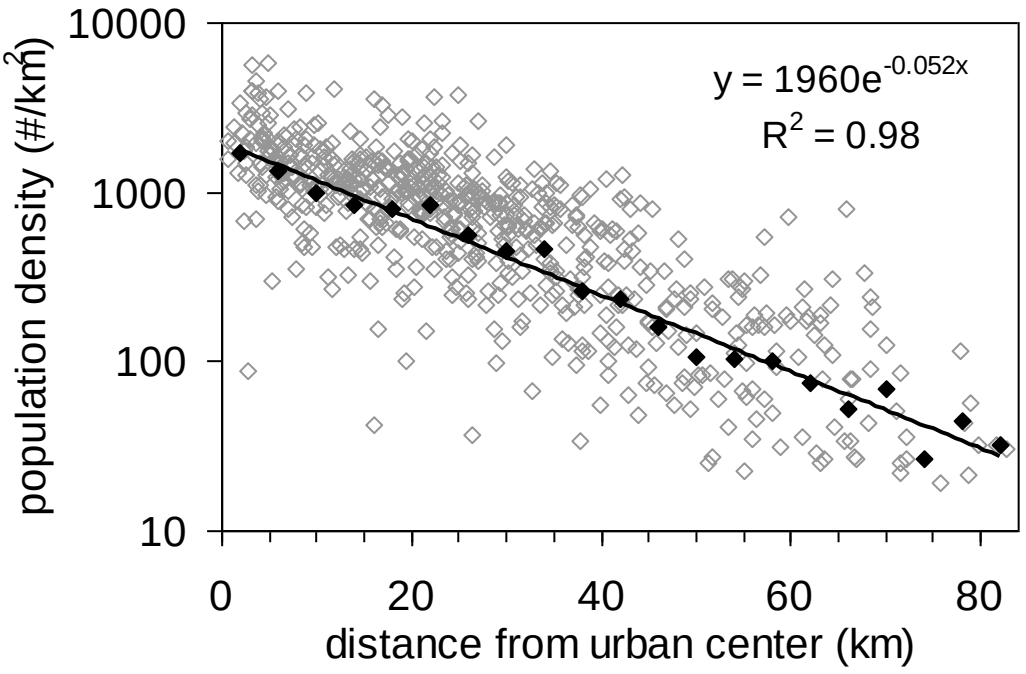


Figure A.1. Population density versus distance from Site A. Population density for 660 census tracts are plotted (grey) and averaged in 4 km increments (black). Least squares regression was performed on the population incremental averages. The regressed relationship between population density and distance from Site A was used to compute the integrated population-weighted semivariance values.

Table A.1. Modified semivariogram parameters. Partial sill (γ_e') is 1 - γ_o'.

	semivariogram nugget (γ_o')		range, km ($3a_e$)		pop-wt semivariance ($\overline{\gamma'}$)	
	concentration	log concentration	concentration	log concentration	concentration	log concentration
1-hr max NO_2	0.040	0.037	153	176	0.461	0.422
1-hr max NO_x	0.031	0.023	254	266	0.326	0.311
8-hr max O_3	0.008	0.010	2,736	3,021	0.043	0.042
1-hr max SO_2	0.053	0.053	98	180	0.596	0.425
1-hr max CO	0.036	0.025	231	243	0.352	0.333
24-hr PM_{10}	0.053	0.048	731	765	0.170	0.161
24-hr $PM_{2.5}$	0.029	0.037	1,873	1,870	0.079	0.086
24-hr $PM_{2.5}$-SO_4	0.023	0.017	2,325	2,375	0.063	0.057
24-hr $PM_{2.5}$-NO_3	0.013	0.033	1,383	1,067	0.080	0.117
24-hr $PM_{2.5}$-NH_4	0.043	0.069	1,071	1,415	0.126	0.131
24-hr $PM_{2.5}$-EC	0.047	0.051	251	331	0.340	0.285
24-hr $PM_{2.5}$-OC	0.082	0.093	975	1,327	0.169	0.157

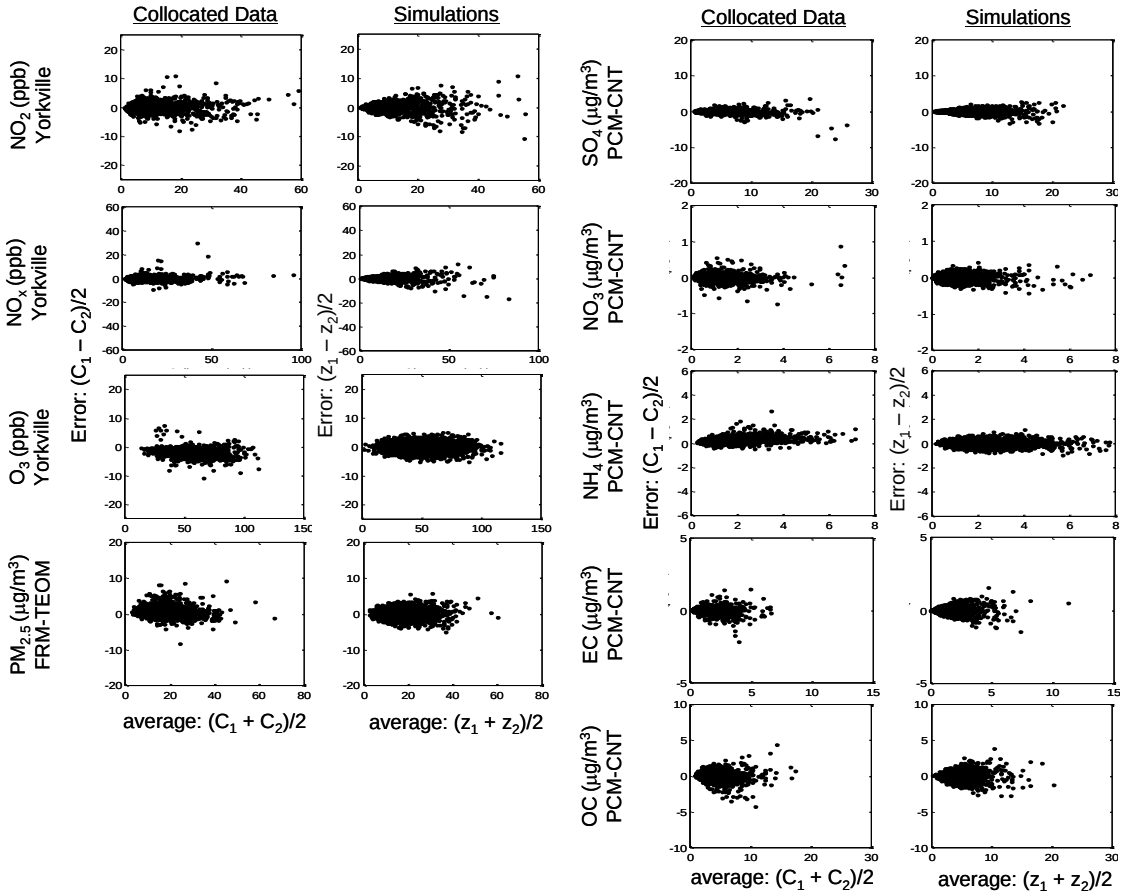
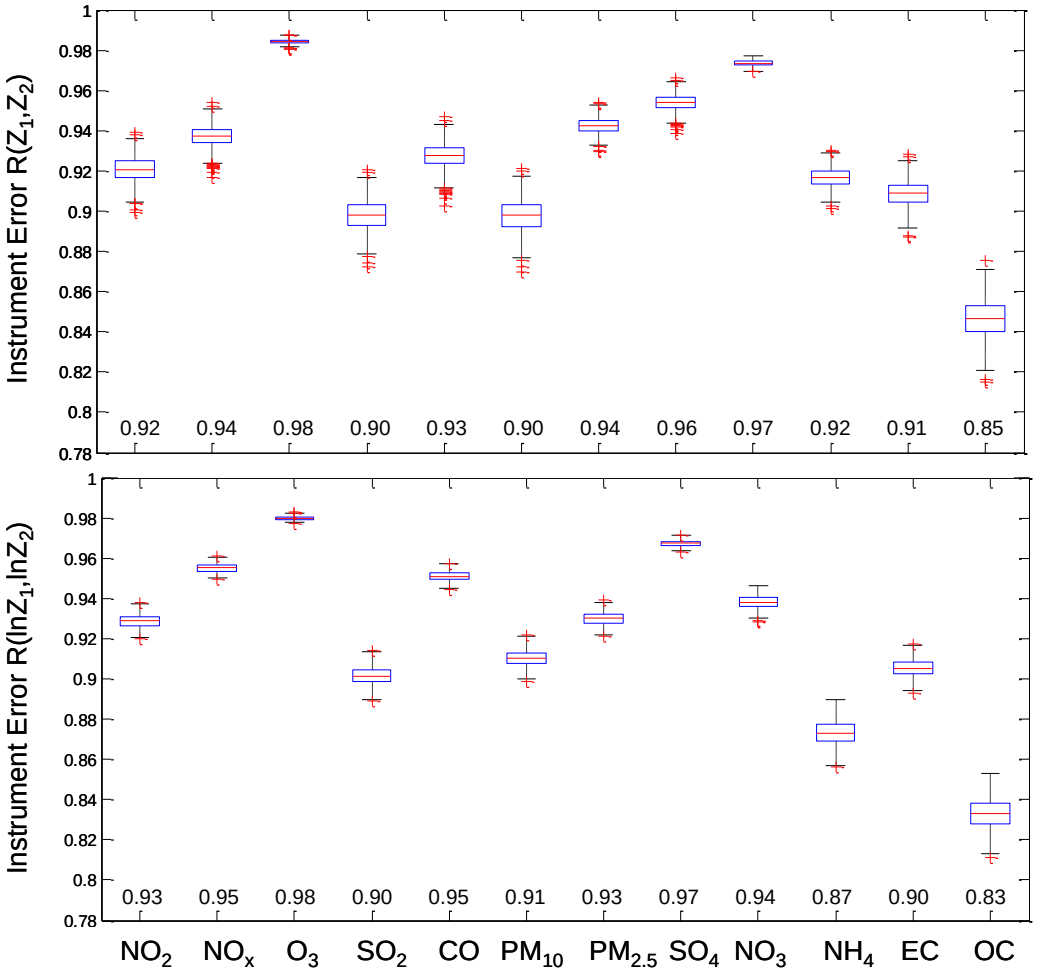


Figure A.2. Collocated instrument precision error and simulated error. Error is calculated as half the difference between collocated measurements (left panel) and between simulations (right panel) for nine pollutants. Bias in continuous measures (TEOM and CNT) has been removed.

Table A.2. Error model optimization parameters and correlation between simulation (*Z)* and base case (Z^*).

		a	***b***	**R (Z, Z*)**
instrument precision error	1-hr max NO_2	0.77 ± 0.12	-0.28 ± 0.04	0.960 ± 0.002
	1-hr max NO_x	0.20 ± 0.02	-0.03 ± 0.03	0.969 ± 0.002
	8-hr max O_3	1.05 ± 0.09	-0.62 ± 0.02	0.992 ± 0.0004
	1-hr max SO_2	0.59 ± 0.03	-0.26 ± 0.02	0.948 ± 0.003
	1-hr max CO	0.23 ± 0.01	0.09 ± 0.03	0.963 ± 0.004
	24-hr PM_{10}	0.55 ± 0.07	-0.17 ± 0.03	0.948 ± 0.003
	24-hr $PM_{2.5}$	1.24 ± 0.15	-0.58 ± 0.04	0.971 ± 0.002
	24-hr $PM_{2.5}$-SO_4	0.18 ± 0.01	0.02 ± 0.03	0.977 ± 0.002
	24-hr $PM_{2.5}$-NO_3	0.18 ± 0.01	-0.55 ± 0.03	0.987 ± 0.001
	24-hr $PM_{2.5}$-NH_4	0.53 ± 0.02	-0.70 ± 0.04	0.957 ± 0.002
	24-hr $PM_{2.5}$-EC	0.33 ± 0.01	-0.18 ± 0.03	0.953 ± 0.003
	24-hr $PM_{2.5}$-OC	0.60 ± 0.05	-0.22 ± 0.05	0.920 ± 0.005
error due to spatial variability	1-hr max NO_2	2.28 ± 0.14	0.03 ± 0.09	0.361 ± 0.034
	1-hr max NO_x	1.43	-0.28	0.499 ± 0.031
	8-hr max O_3	0.42 ± 0.01	-0.09 ± 0.03	0.916 ± 0.007
	1-hr max SO_2	2.37 ± 0.17	-0.07 ± 0.09	0.325 ± 0.034
	1-hr max CO	1.69 ± 0.11	0.02 ± 0.19	0.462 ± 0.032
	24-hr PM_{10}	0.93 ± 0.03	-0.12 ± 0.06	0.709 ± 0.020
	24-hr $PM_{2.5}$	0.62 ± 0.02	-0.18 ± 0.03	0.853 ± 0.011
	24-hr $PM_{2.5}$-SO_4	0.50 ± 0.01	-0.05 ± 0.04	0.879 ± 0.010
	24-hr $PM_{2.5}$-NO_3	0.66 ± 0.02	-0.47 ± 0.04	0.850 ± 0.011
	24-hr $PM_{2.5}$-NH_4	0.80 ± 0.03	-0.19 ± 0.04	0.774 ± 0.016
	24-hr $PM_{2.5}$-EC	1.12 ± 0.18	-0.67 ± 0.21	0.489 ± 0.053
	24-hr $PM_{2.5}$-OC	0.93 ± 0.03	-0.10 ± 0.06	0.709 ± 0.020


Instrument Error R(Z1,Z2)
1
0.98
0.96
0.94
0.92
0.9
0.88
0.86
0.84
0.82
0.8
0.78
0.92 0.94 0.98 0.90 0.93 0.90 0.94 0.96 0.97 0.92 0.91 0.85
Instrument Error R(lnZ1,lnZ2)
0.93 0.95 0.98 0.90 0.95 0.91 0.93 0.97 0.94 0.87 0.90 0.83
NO2 NOx O3 SO2 CO PM10 PM2.5 SO4 NO3 NH4 EC OC

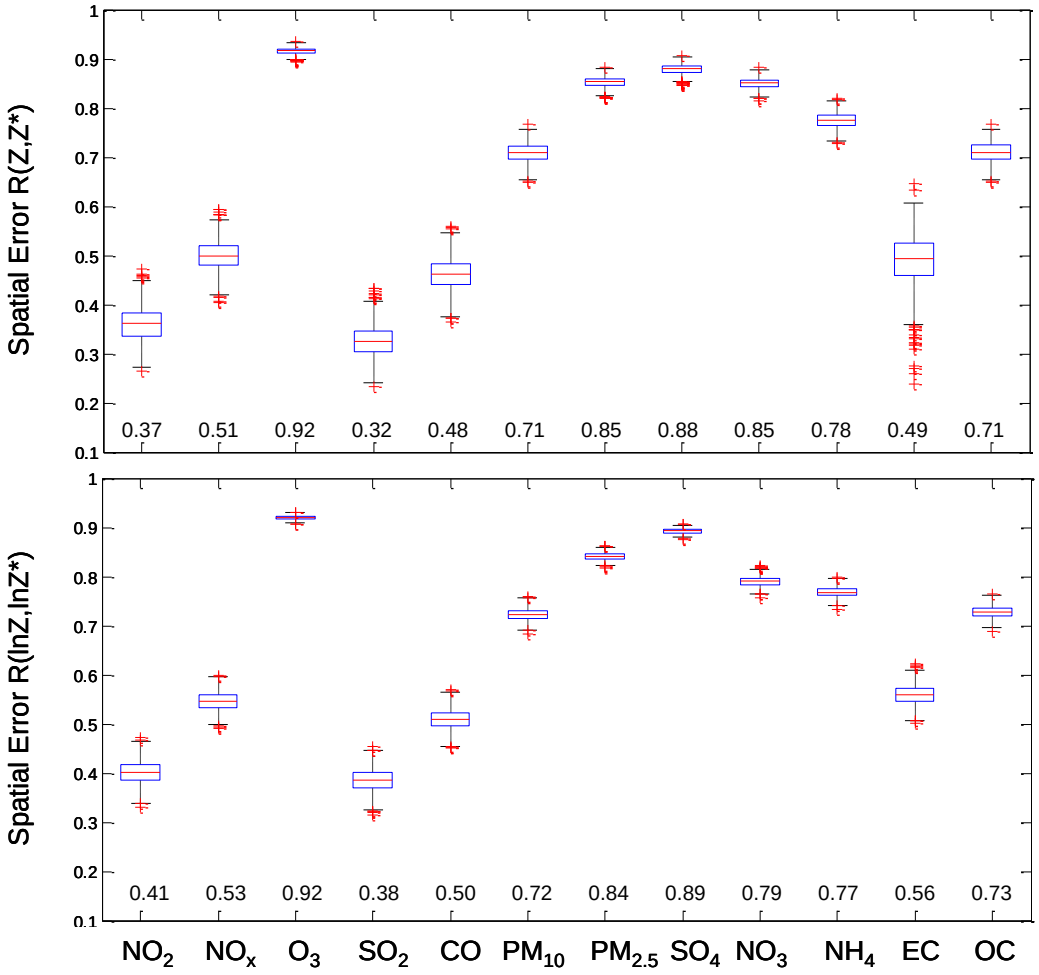


Figure A.3. Boxplots of the distribution of correlation coefficients between 1000 simulations for instrument error and between 1000 simulations and the base case time-series for spatial error. Target values shown at bottom.

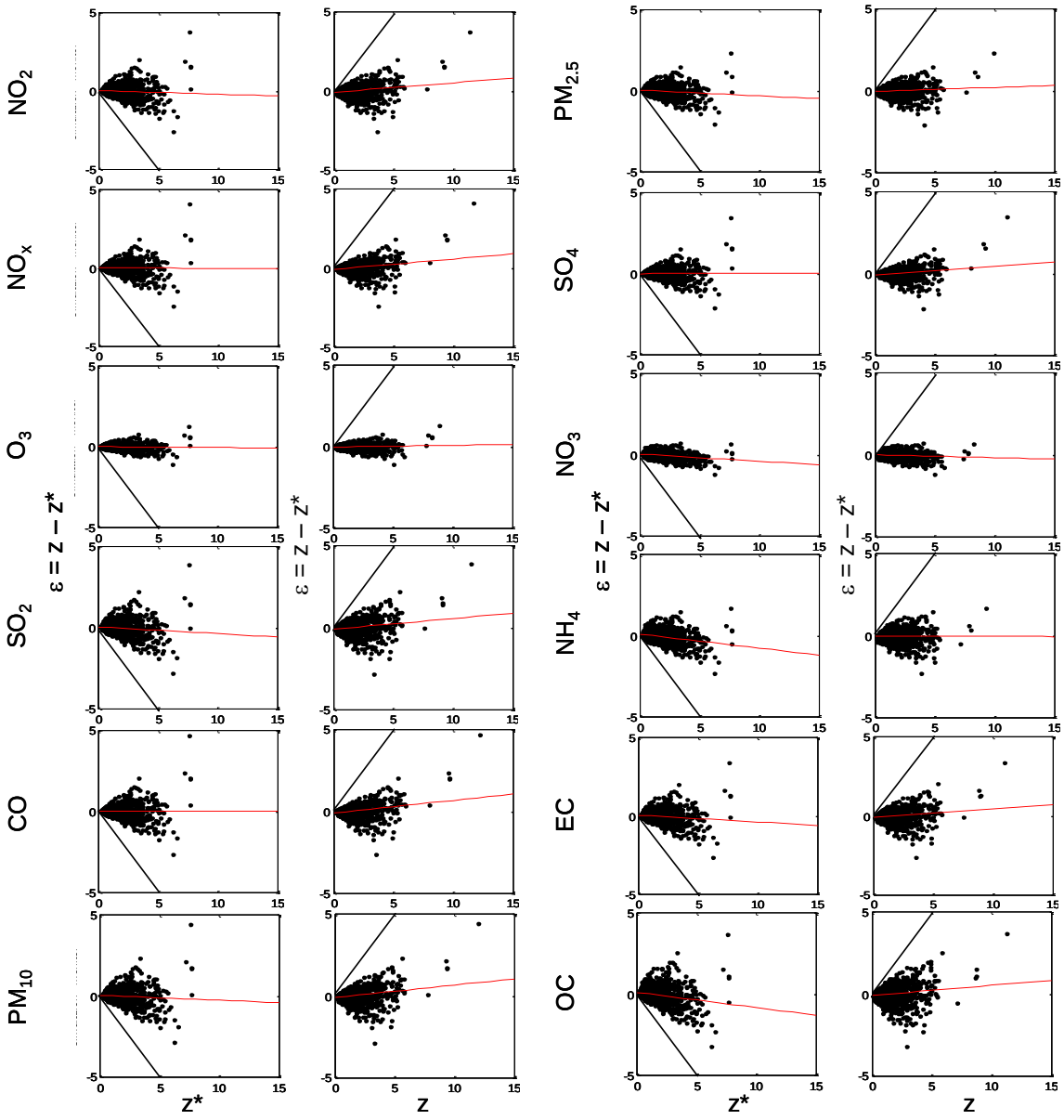


Figure A.4. Scatterplots of error versus the base case (Z^*) and versus the simulation (Z) for a sample Monte Carlo simulation of instrument error. When error is independent of the true value (here, defined as the base case monitor data), the error type is classical. When error is independent of the measurement (here, defined as the simulation which has error added), the error type is Berkson. The above plots suggest that this error is neither classical nor Berkson.

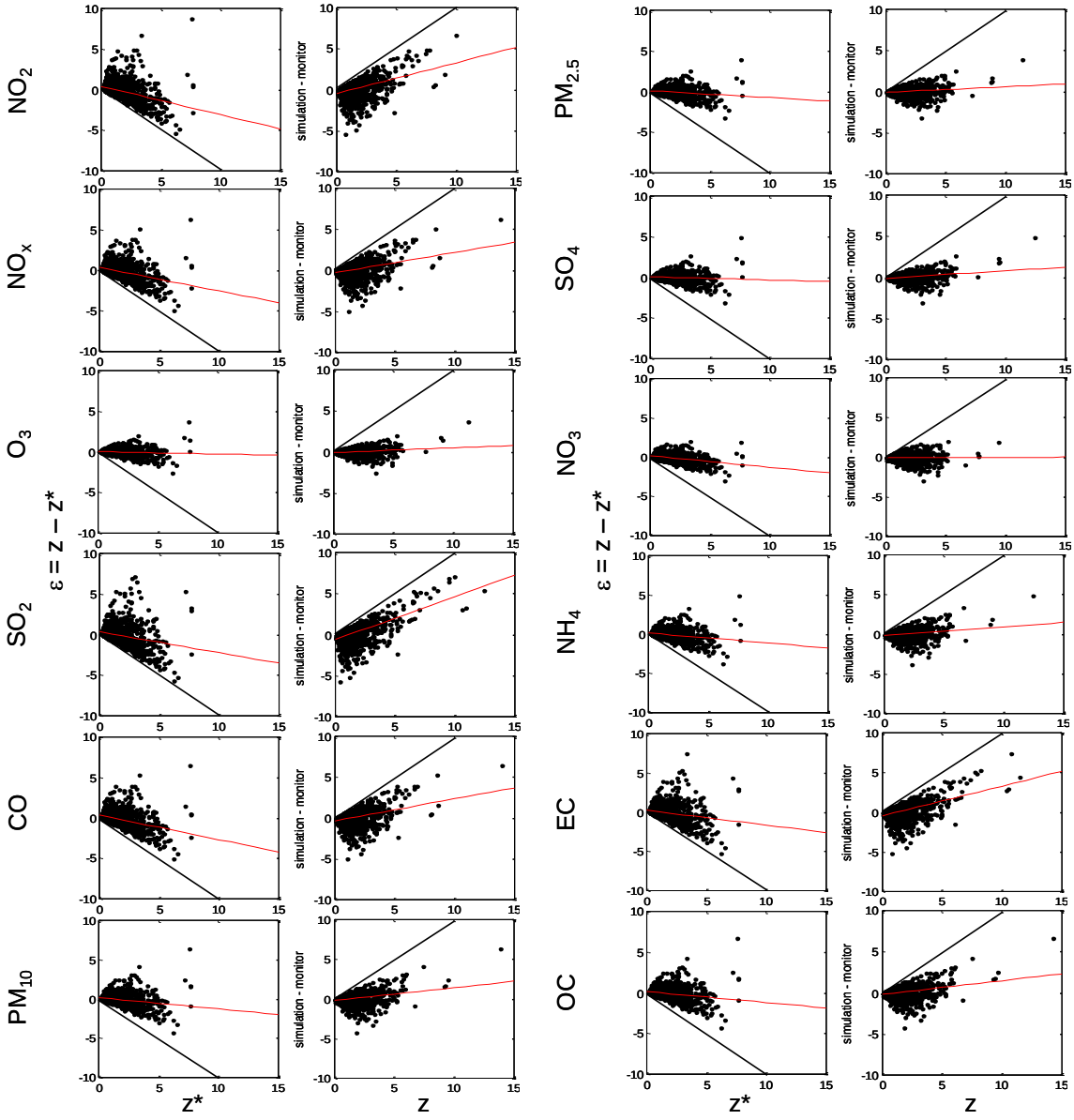


Figure A.5. Scatterplots of error versus the base case (Z^*) and versus the simulation (Z) for a sample Monte Carlo simulation of spatial error. When error is independent of the true value (here, defined as the base case monitor data), the error type is classical. When error is independent of the measurement (here, defined as the simulation which has error added), the error type is Berkson. The above plots suggest that this error is neither classical nor Berkson.

APPENDIX B

SUPPLEMENTAL MATERIAL FOR ERROR TYPE ASSESSMENT

Table B.1. Power transformation analysis. The Hinkley d_λ statistic, defined as $\frac{|mean - median|}{IQR}$, provides information on the normality of a transformation for a given choice of λ, where $\lambda = 0$ represents a log transformation and $\lambda = 1$ represents untransformed data (Hinkley 1977). The smaller d_λ, the closer the data transformation is to a normal distribution. Hinkley d_λ statistics are shown for daily measurements at the central monitoring site over the period 1999-2004.

	d_λ Measurement Data	
Pollutant	$\lambda = 0$	$\lambda = 1$
1-hr max NO_2	0.069	0.079
1-hr max NO_x	0.058	0.376
8-hr max O_3	0.121	0.095
1-hr max SO_2	0.050	0.271
1-hr max CO	0.080	0.388
24-hr PM_{10}	0.018	0.140
24-hr $PM_{2.5}$	0.038	0.134
24-hr $PM_{2.5}$-SO_4	0.0009	0.252
24-hr $PM_{2.5}$-NO_3	0.043	0.333
24-hr $PM_{2.5}$-NH_4	0.024	0.179
24-hr $PM_{2.5}$-EC	0.019	0.233
24-hr $PM_{2.5}$-OC	0.002	0.204

$$T_1(x) = \begin{cases} x^\lambda, & \lambda > 0 \\ \ln(x), & \lambda = 0 \end{cases}$$

where *x* is a daily measurement of pollutant concentration

Derivations of equations in text for error models.

Z^* = true ambient, log-normally distributed; given
Z = measured ambient, log-normally distributed; simulated

1. Type C error model

$$\chi^* = \frac{\ln Z^* - \mu_{\ln Z^*}}{\sigma_{\ln Z^*}} \qquad \sim N(0,1) \qquad \text{(B.1)}$$

$$\chi = \chi^* + \varepsilon_\chi \qquad \sim N(0, \sigma_\chi = \sqrt{1+\sigma_{err}^{\ 2}}) \qquad \text{(B.2)}$$

where $\varepsilon_\chi = N\ \sigma_{err}$ $\sim N(0, \sigma_{err})$

N = random number $\sim N(0,1)$

$$\ln Z = \chi\sigma_{\ln Z^*} + \mu_{\ln Z^*} \qquad \sim N(\mu_{\ln Z^*}, \sigma_{\ln Z} = \sigma_{\ln Z^*}\sqrt{1+\sigma_{err}^{\ 2}}) \qquad \text{(B.3)}$$

2. Type B error model

$$\chi^* = \frac{\ln Z^* - \mu_{\ln Z^*}}{\sigma_{\ln Z^*}} = \chi + \varepsilon_\chi \qquad \sim N(0,1) \qquad \text{(B.4)}$$

$$\chi = \frac{\chi^* + \varepsilon_\chi}{1+\sigma_{err}^{\ 2}} \qquad \sim N\left(0, \sigma_\chi = \frac{1}{\sqrt{1+\sigma_{err}^{\ 2}}}\right) \qquad \text{(B.5)}$$

$$\ln Z = \chi\sigma_{\ln Z^*} + \mu_{\ln Z^*} \qquad \sim N\left(\mu_{\ln Z^*}, \sigma_{\ln Z} = \frac{\sigma_{\ln Z^*}}{\sqrt{1+\sigma_{err}^{\ 2}}}\right) \qquad \text{(B.6)}$$

3. Relationship between Pearson correlation coefficient and amount of error added (σ_{err})

$$\bar{R}(\chi, \chi^*) = \sqrt{\bar{R}(\chi_i, \chi_j)} = \frac{1}{\sqrt{1+\sigma_{err}^{\ 2}}} \qquad \text{(B.7)}$$

$$\sigma_{err} = \sqrt{\frac{1-\bar{R}(\chi_i, \chi_j)}{\bar{R}(\chi_i, \chi_j)}} \qquad \text{(B.8)}$$

4. Population-weighted semivariance and Pearson correlation coefficient

$$\text{semivariance: } \bar{\gamma}'(\chi_i, \chi_j) = \frac{1}{p_{total}}\left(\sum_i^{660}\sum_{j=i+1}^{660} p_{i,j}\gamma'_{i,j}(h) + \sum_i^{660} p_i\gamma'_i(r)\right) \qquad \text{(B.9)}$$

where $\gamma'_{i,j}$ = inter-tract semivariance, based on distance h between tract i and j centroids

γ'_i = intra-tract semivariance, estimated from average radius *r* of census tract *i*
p_{total} = total number of pairs of residences = ½ ($n_{total}^2 - n_{total}$)
$p_{i,j}$ = ½ (n_i x n_j) ; p_i = ½ ($n_i^2 - n_i$)
n_i = population of tract *i*; n_{total} = total population

Pearson correlation coefficient: $\overline{R}(\chi_i, \chi_j) = \dfrac{1 - \overline{\gamma}'(\chi_i, \chi_j)}{1 + \overline{\gamma}'(\chi_i, \chi_j)}$ (B.10)

APPENDIX C

SUPPLEMENTAL MATERIAL FOR TRUE AMBIENT FIELD SIMULATION

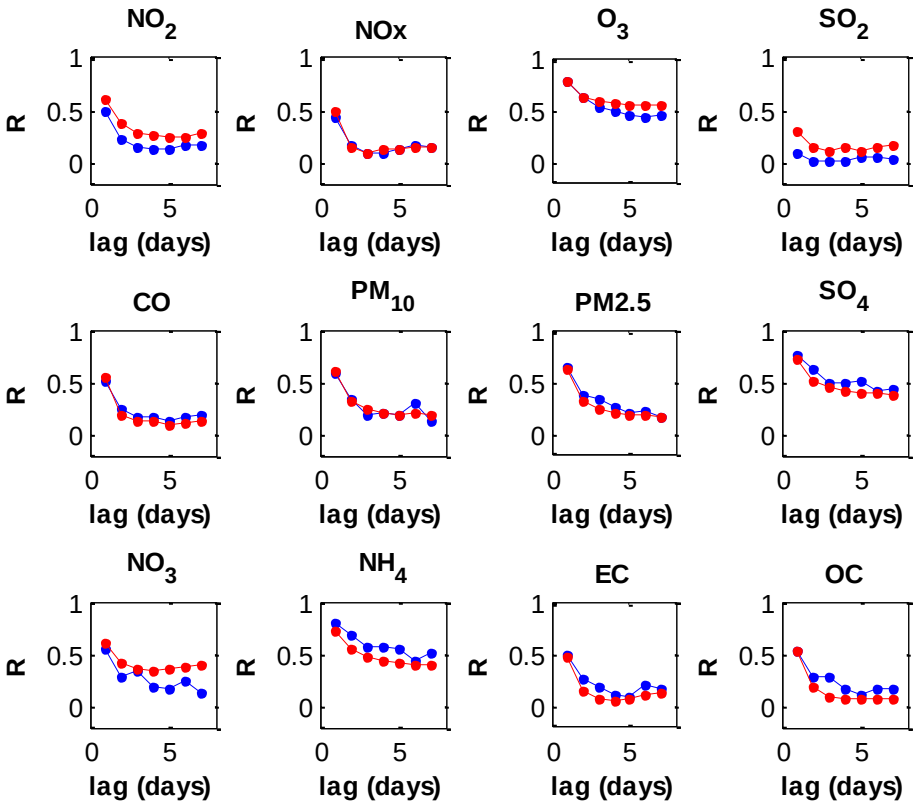


Figure C.1. Short-term temporal autocorrelation of measurements at the central monitoring site (red) and of the simulated time-series at the central grid cell (blue).

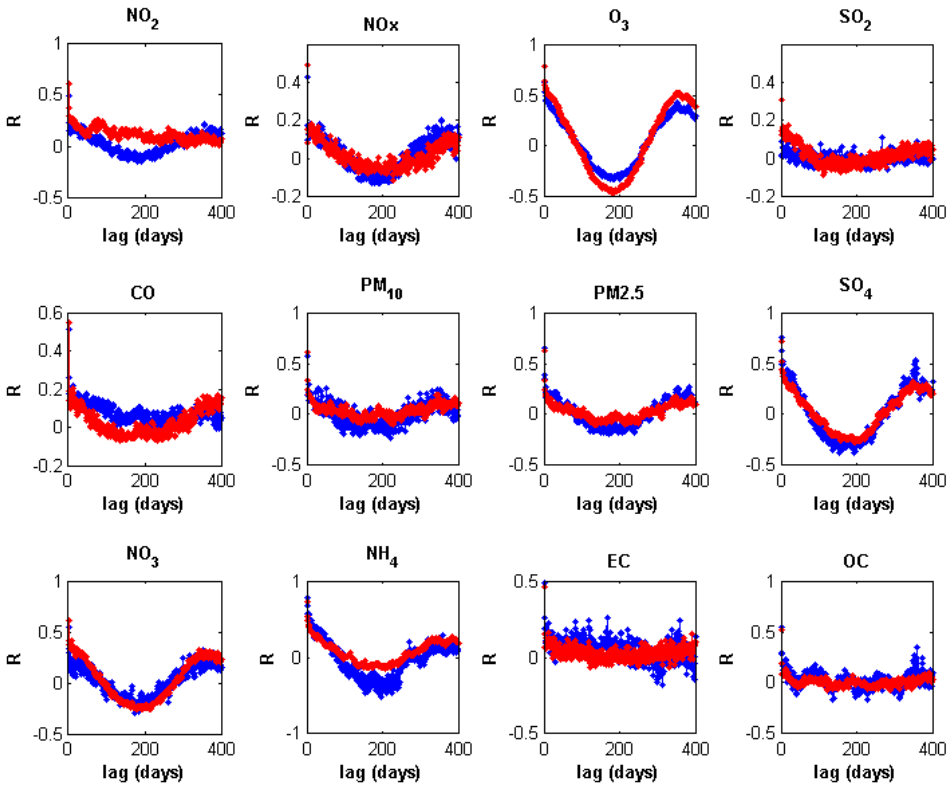


Figure C.2. Long term temporal autocorrelation of measurements at the urban (red) and rural (blue) monitoring sites.

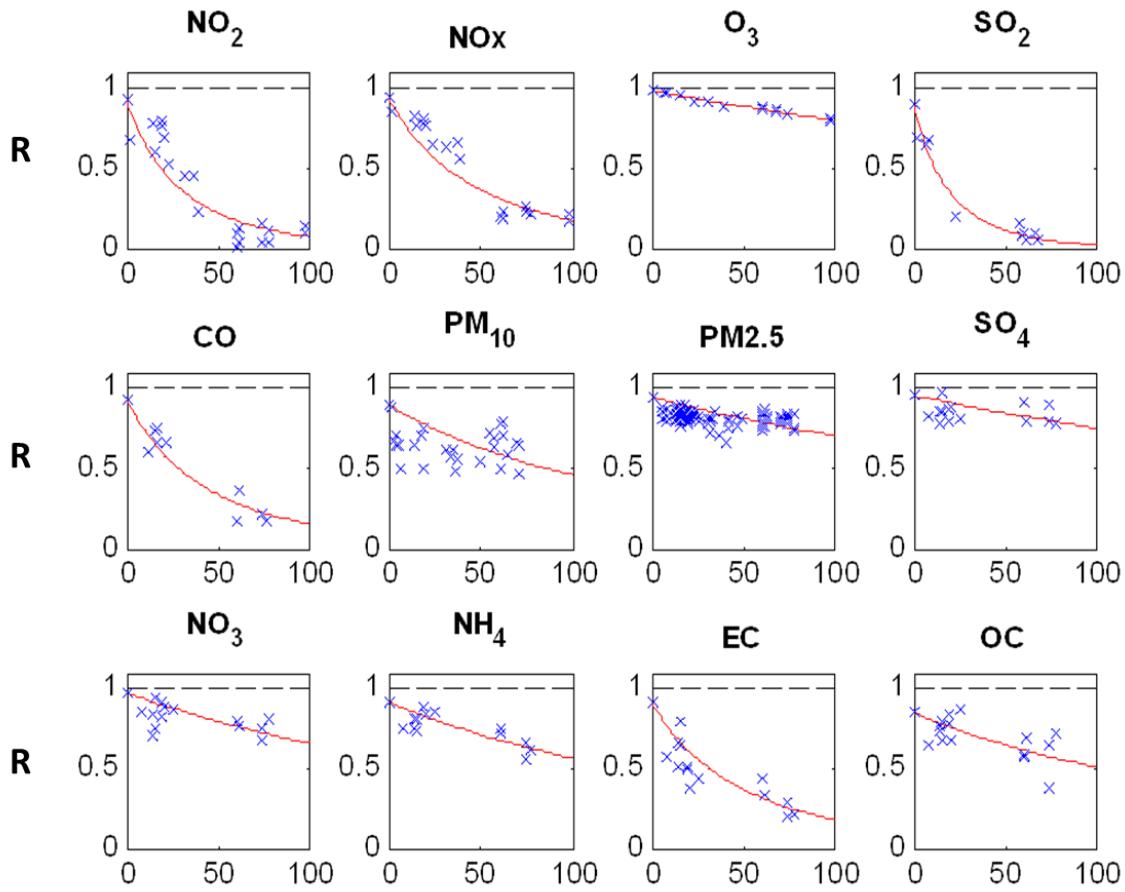


Figure C.3. Spatial correlograms between monitoring site time-series. Plotted points represent monitor pairs. The curve is an exponential regression of the plotted points.

Table C.1. Population-weighted average spatial autocorrelations.

pollutant	R
NO_2	0.318
NO_x	0.454
O_3	0.901
SO_2	0.208
CO	0.426
PM_{10}	0.672
$PM_{2.5}$	0.832
SO_4	0.863
NO_3	0.822
NH_4	0.744
EC	0.453
OC	0.680

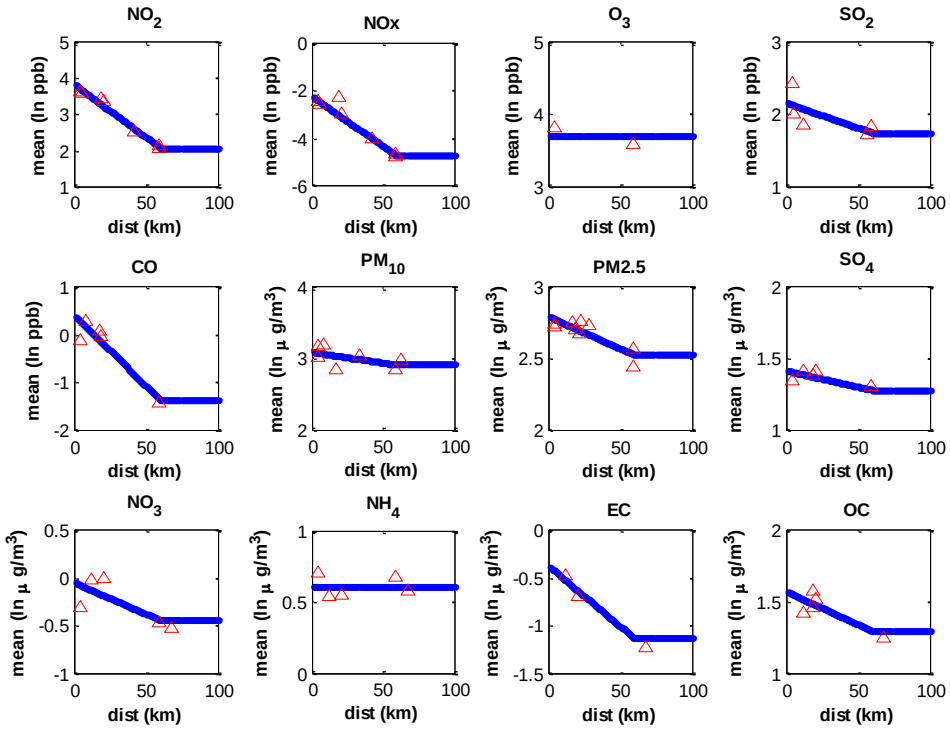


Figure C.4. Plotted logged monitoring site data time-series mean (red triangles) versus distance from urban center and regressed linear functions up to 60 km with fixed rural background levels beyond 60 km (blue).

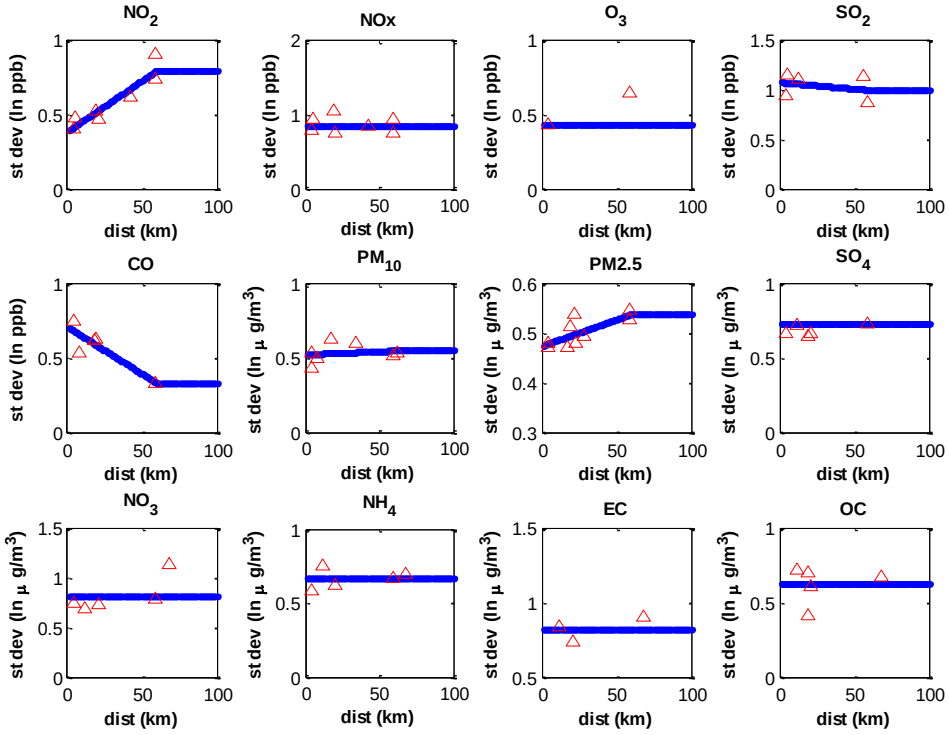


Figure C.5. Plotted logged monitoring site data time-series standard deviation (red triangles) versus distance from urban center and regressed linear functions up to 60 km with fixed rural background levels beyond 60 km (blue).

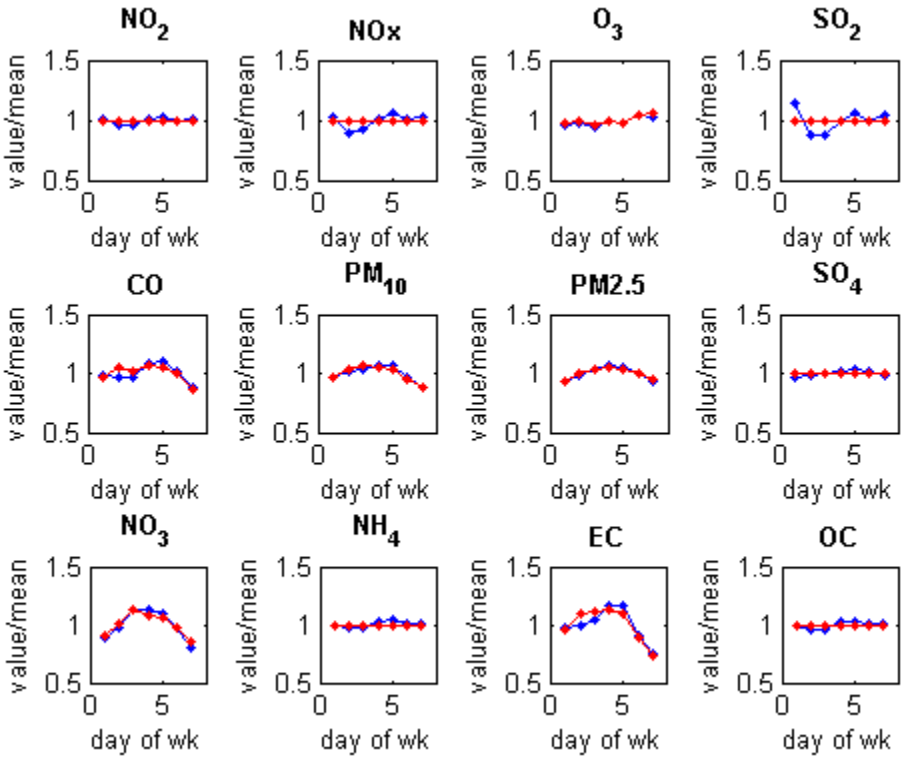
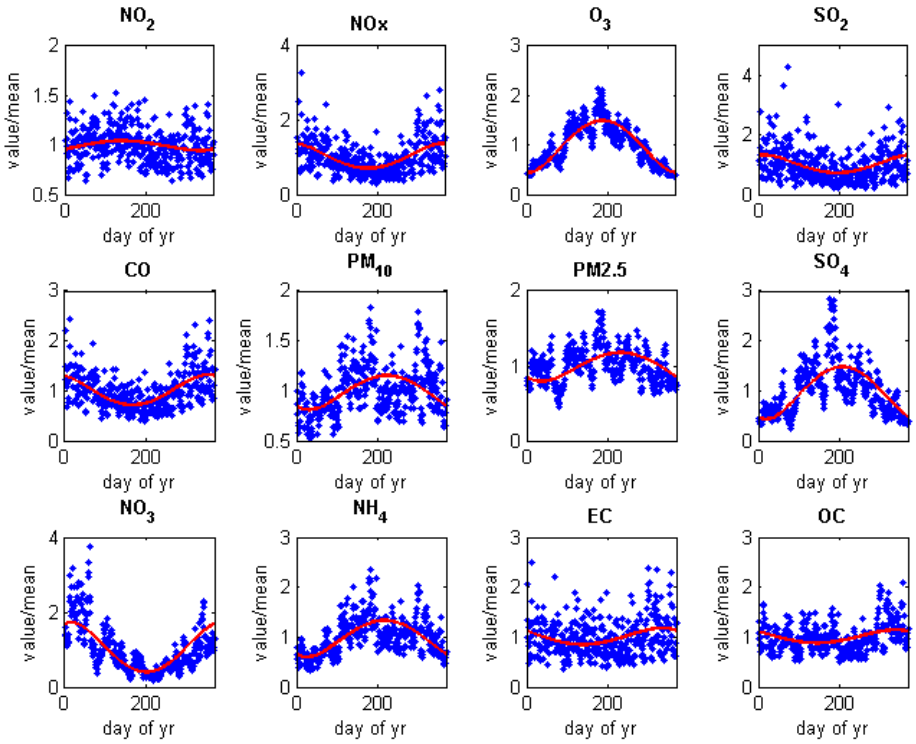


Figure C.6. Concentration normalized by time-series mean for each day of week (Monday = 1, Sunday = 7) for central monitor data (red) and simulated time-series (blue).

(a)



(b)

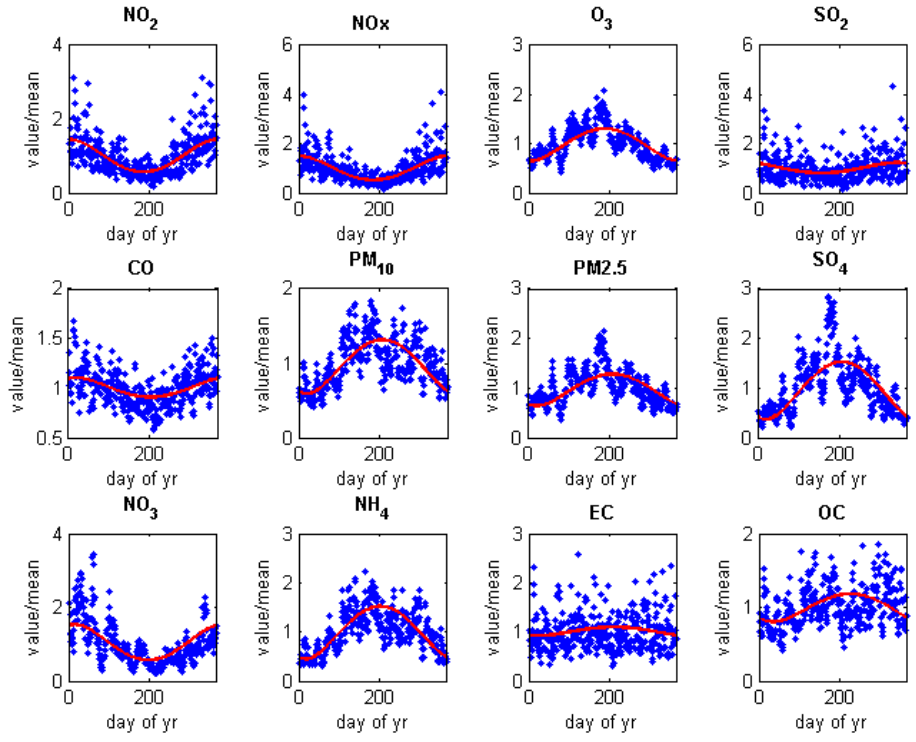


Figure C.7. Concentration normalized by time-series mean for each day of year (1-365) for simulated time-series at urban location (a) and rural location (b). Red curves are fourth order regressions fit to monitor data.

Table C.2. Mean and standard deviation linear model parameters: slope, *m*, and y-intercept, *b*, used for simulation of true ambient concentration field.

	μ_r		σ_r	
pollutant	*m*	*b*	*m*	*b*
NO_2	-0.029	3.839	0.007	0.391
NO_x	-0.042	-2.248	0	0.845
O_3	0	3.695	0	0.434
SO_2	-0.007	2.162	-0.001	1.080
CO	-0.029	0.388	-0.006	0.710
PM_{10}	-0.003	3.094	0	0.527
$PM_{2.5}$	-0.004	2.793	0.001	0.475
SO_4	-0.002	1.417	0	0.733
NO_3	-0.006	-0.049	0	0.823
NH_4	0	0.608	0	0.665
EC	-0.013	-0.374	0	0.828
OC	-0.005	1.573	0	0.630

Table C.3. Fourth order polynomial regression coefficients ($ax^4+bx^3+cx^2+dx+e$) where x is day of year (1-365) for seasonal trend functions at the urban center

pollutant	*a*	*b*	*c*	*d*	*e*
NO_2	6.64E-11	-3.5E-08	1.48E-06	0.000896	0.956689
NO_x	-5.9E-10	4.13E-07	-6.8E-05	-0.00135	1.359902
O_3	9.45E-10	-7E-07	0.000129	-0.00041	0.442271
SO_2	-5.4E-10	4.13E-07	-8.3E-05	0.001395	1.32044
CO	-5.2E-10	3.41E-07	-4.8E-05	-0.00254	1.303274
PM_{10}	2.5E-10	-2.2E-07	5.59E-05	-0.00275	0.847483
$PM_{2.5}$	2.54E-10	-2.4E-07	6.35E-05	-0.00361	0.849826
SO_4	8.74E-10	-7.2E-07	0.00016	-0.00529	0.478189
NO_3	-1.1E-09	9.1E-07	-0.00019	0.00482	1.690008
NH_4	5.7E-10	-4.9E-07	0.000117	-0.00503	0.659703
EC	-2.2E-10	1.14E-07	-3.7E-06	-0.00311	1.128679
OC	-1.9E-10	1.04E-07	-6.7E-06	-0.00225	1.111637

Table C.4. Fourth order polynomial regression coefficients ($ax^4+bx^3+cx^2+dx+e$) where x is day of year (1-365) for seasonal trend functions at a rural location (60 km from urban center).

pollutant	*a*	*b*	*c*	*d*	*e*
NO_2	-7.95E-10	5.81E-07	-0.00011	5.28E-05	1.445633
NO_x	-8.77E-10	6.4E-07	-0.00012	-4.8E-05	1.492518
O_3	5.86E-10	-4.4E-07	8.43E-05	-0.00076	0.653412
SO_2	-3.39E-10	2.07E-07	-2.3E-05	-0.00266	1.197628
CO	-1.67E-10	1.35E-07	-2.9E-05	0.000843	1.097222
PM_{10}	6.061E-10	-5E-07	0.000111	-0.00368	0.62078
$PM_{2.5}$	5.256E-10	-4.3E-07	9.78E-05	-0.00339	0.676142
SO_4	9.719E-10	-8E-07	0.000177	-0.00575	0.417184
NO_3	-8.5E-10	6.64E-07	-0.00014	0.002941	1.518032
NH_4	9.182E-10	-7.4E-07	0.000158	-0.00433	0.480039
EC	1.42E-10	-1.2E-07	2.79E-05	-0.0011	0.9309
OC	2.579E-10	-2.4E-07	6.24E-05	-0.00342	0.856541

Table C.5. Day of week model parameter values used for simulation of true ambient concentration field (Monday = 1, Sunday = 7).

	α_{wk}						
pollutant	1	2	3	4	5	6	7
NO_2	1.00	1.00	1.00	1.00	1.00	1.00	1.00
NO_x	1.00	1.00	1.00	1.00	1.00	1.00	1.00
O_3	0.99	1.00	0.96	1.00	0.98	1.05	1.06
SO_2	1.00	1.00	1.00	1.00	1.00	1.00	1.00
CO	0.96	1.05	1.01	1.07	1.05	1.00	0.86
PM_{10}	0.97	1.04	1.06	1.06	1.04	0.95	0.87
$PM_{2.5}$	0.94	0.99	1.04	1.06	1.03	0.99	0.95
SO_4	1.00	1.00	1.00	1.00	1.00	1.00	1.00
NO_3	0.91	1.02	1.13	1.08	1.06	0.97	0.85
NH_4	1.00	1.00	1.00	1.00	1.00	1.00	1.00
EC	0.95	1.09	1.12	1.13	1.09	0.89	0.74
OC	1.00	1.00	1.00	1.00	1.00	1.00	1.00

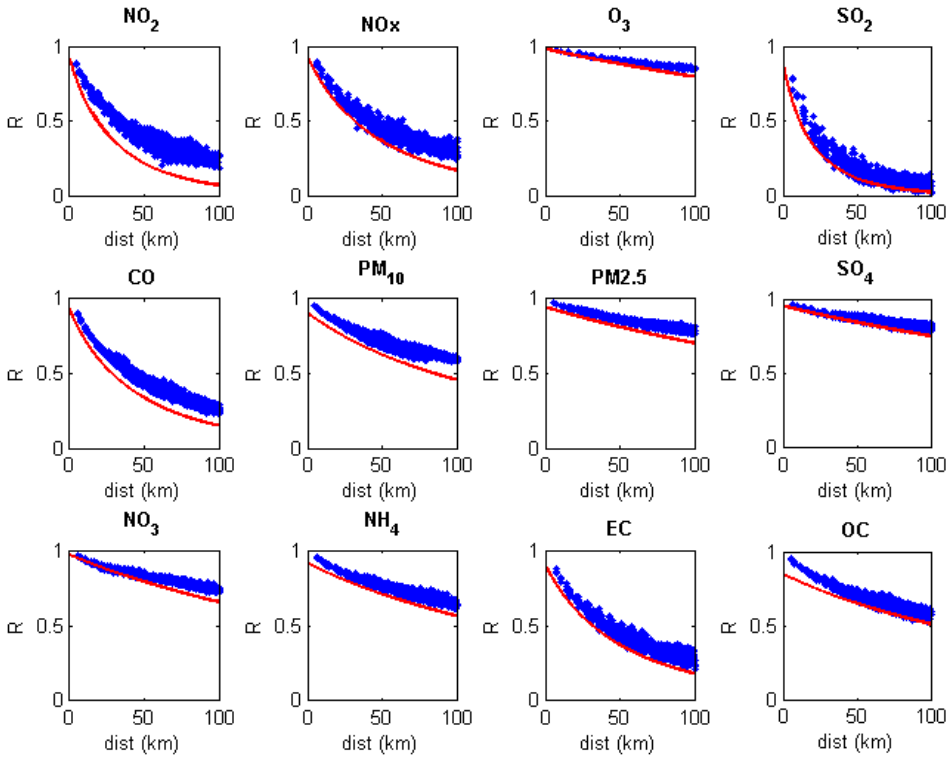
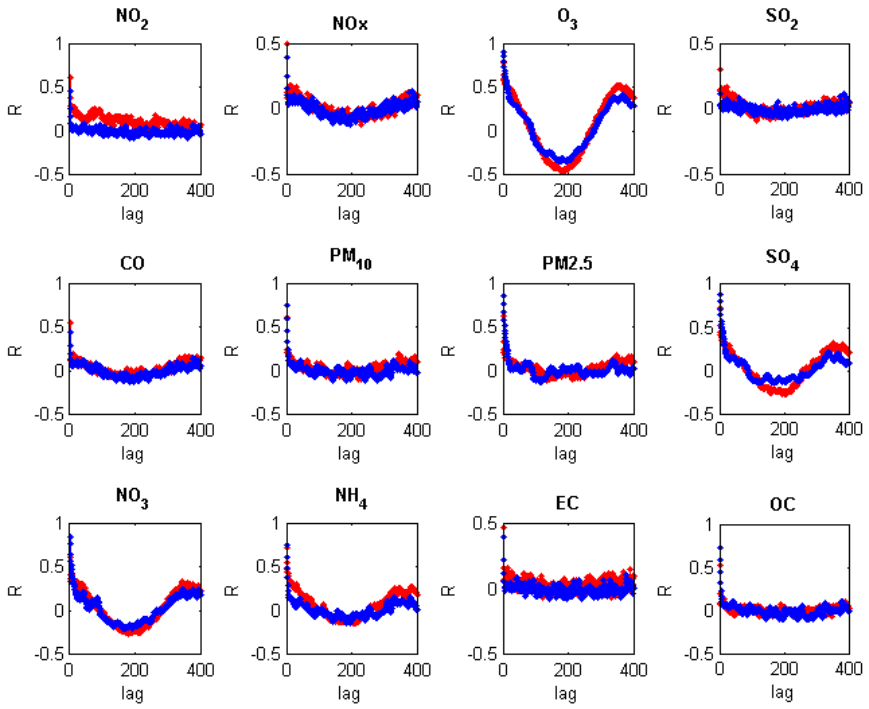


Figure C.8. Spatial correlogram model regressed from monitoring site data (red) and correlation of simulated time-series at all grid cells with the time-series at a rural grid cell (blue).

(a)



(b)

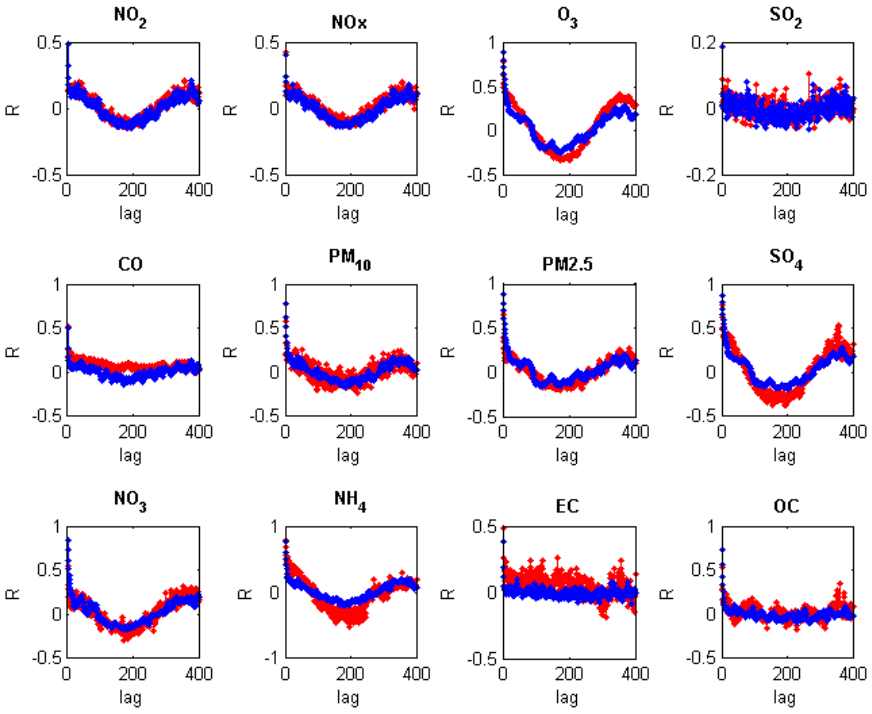


Figure C.9. Long term temporal autocorrelation of measurements (red) and simulated time-series (blue) at urban (a) and rural (b) locations.

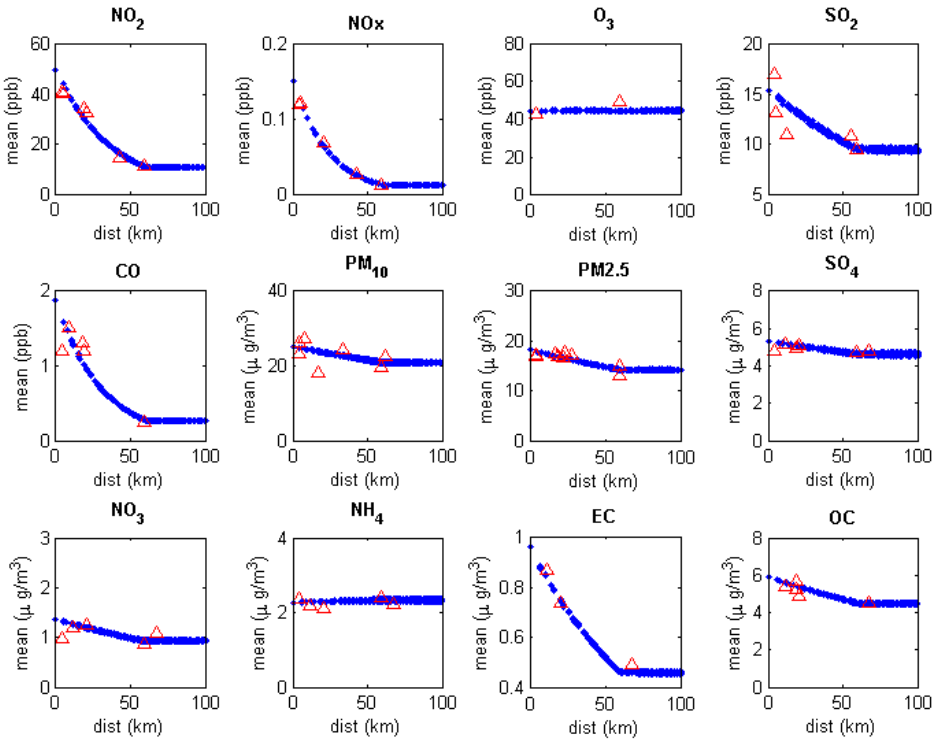


Figure C.10. Time-series mean as a function of distance from urban center. Monitor data shown as red triangles and simulated time-series shown in blue.

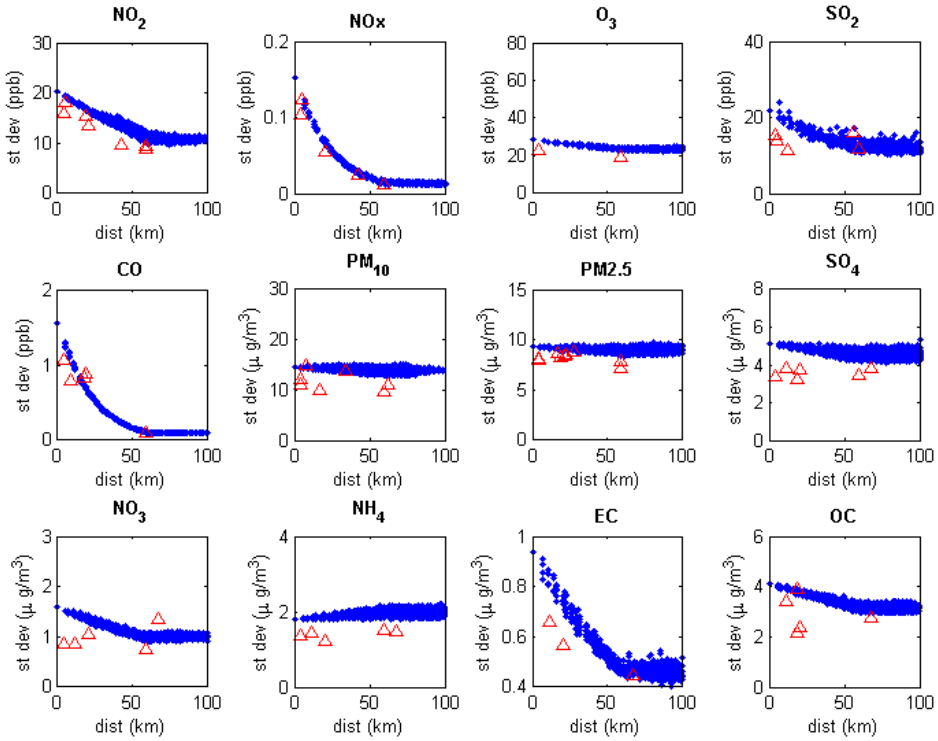
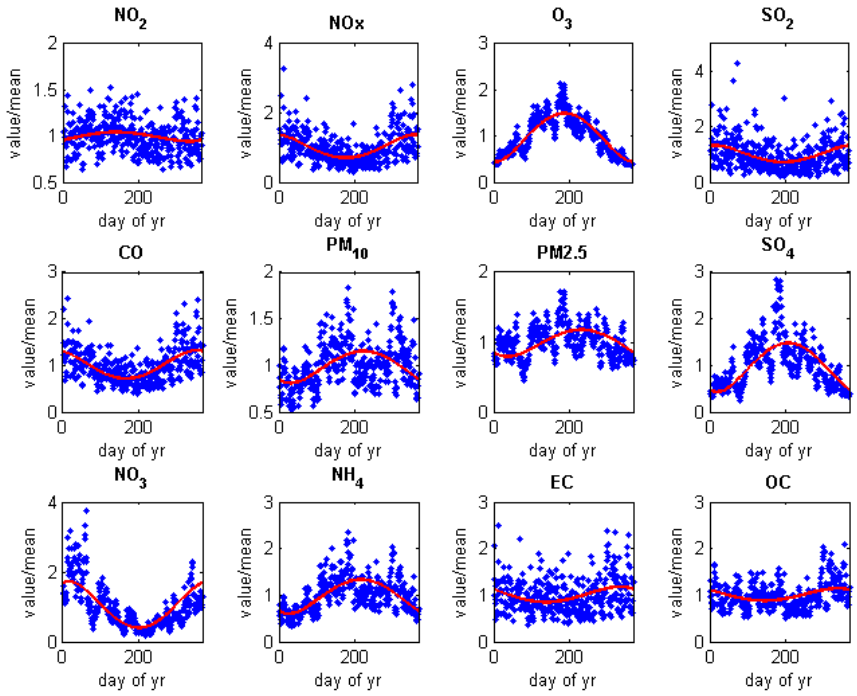


Figure C.11. Time-series standard deviation as a function of distance from urban center. Monitor data shown as red triangles and simulated time-series shown in blue.

(a)



(b)

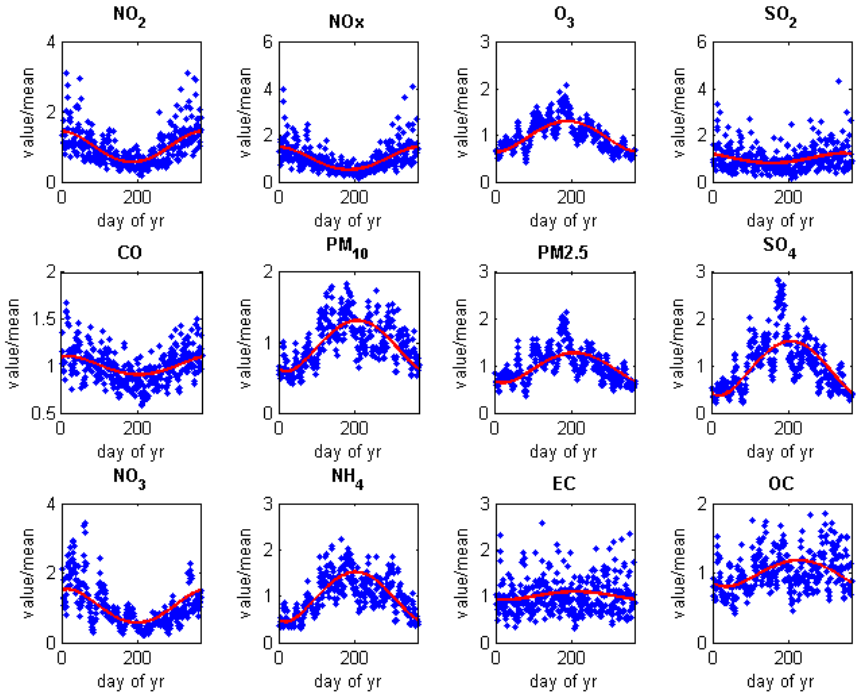


Figure C.12. Concentration normalized by time-series mean for each day of year (1-365) for simulated time-series at urban location (a) and rural location (b). Red curves are 4th order regressions fit to monitor data.

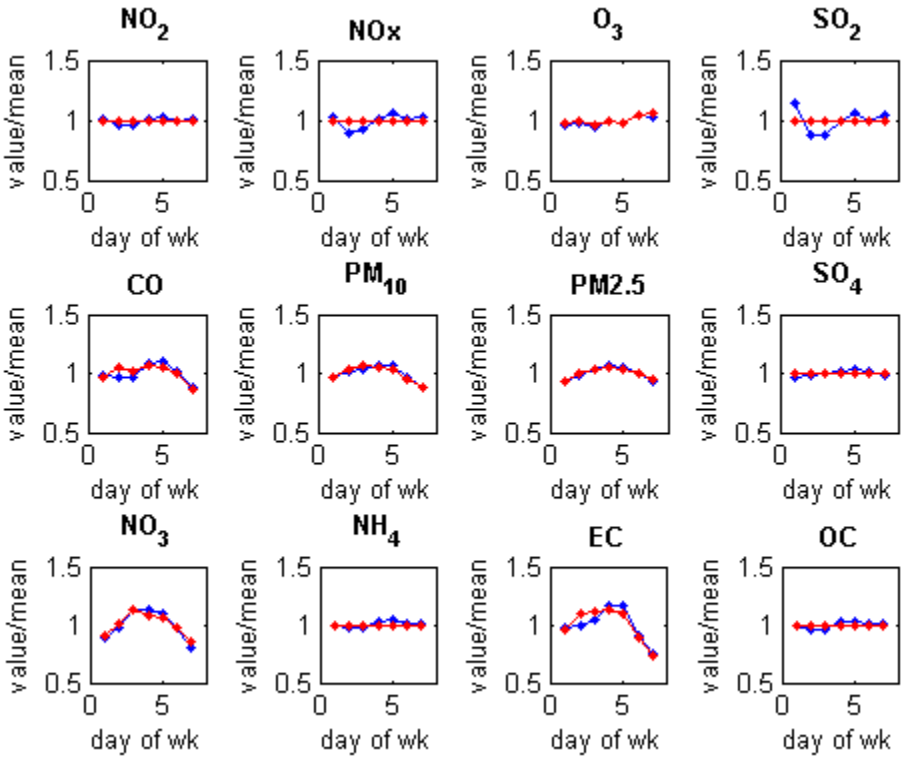


Figure C.13. Concentration normalized by time-series mean for each day of week (Monday = 1, Sunday = 7) for central monitor data (red) and simulated time-series (blue).

Table C.6. Pearson correlation coefficients between the ambient and monitor simulations and the correlation expected from a collocated instrument analysis of a previous study (Goldman et al. 2010).

pollutant	correlation expected from collocated instruments	correlation between ambient and monitor simulations
NO_2	0.959	0.964
NO_x	0.970	0.971
O_3	0.990	0.992
SO_2	0.949	0.945
CO	0.964	0.965
PM_{10}	0.949	0.949
$PM_{2.5}$	0.970	0.971
$PM_{2.5}$-SO_4	0.980	0.949
$PM_{2.5}$-NO_3	0.985	0.988
$PM_{2.5}$-NH_4	0.959	0.963
$PM_{2.5}$-EC	0.954	0.955
$PM_{2.5}$-OC	0.922	0.931

APPENDIX D

SUPPLEMENTAL MATERIAL FOR DALLAS SPATIAL VARIABILITY ASSESSMENT

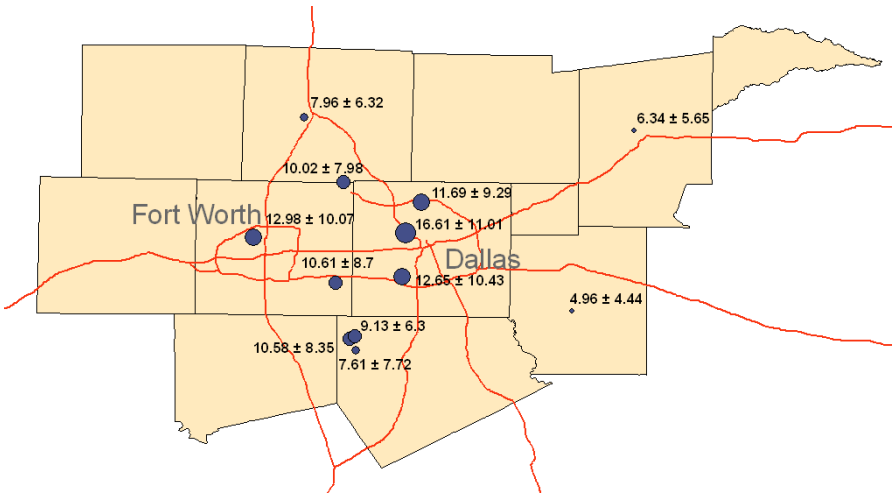


Figure D.1. NO_2 time-series mean (ppb) plotted at monitor locations with graduated symbols for Dallas (a) and Atlanta (b). Labels indicate time-series mean ± time-series standard deviation.

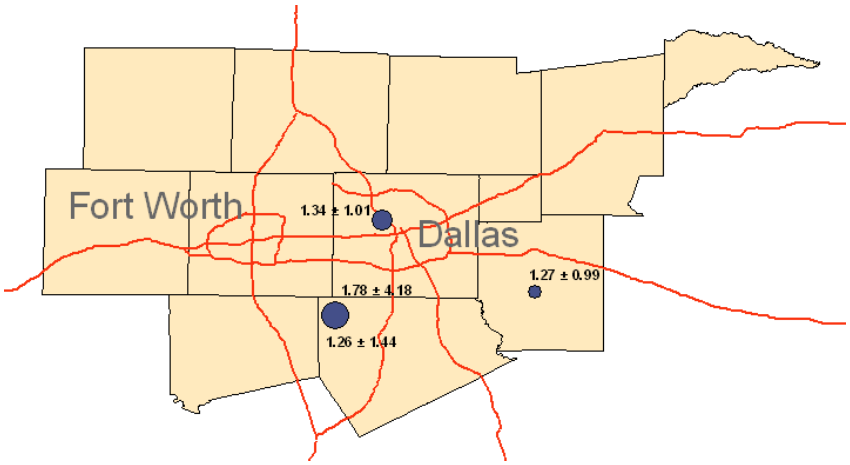


Figure D.2. SO_2 time-series mean (ppb) plotted at monitor locations with graduated symbols for Dallas (a) and Atlanta (b). Labels indicate time-series mean ± time-series standard deviation.

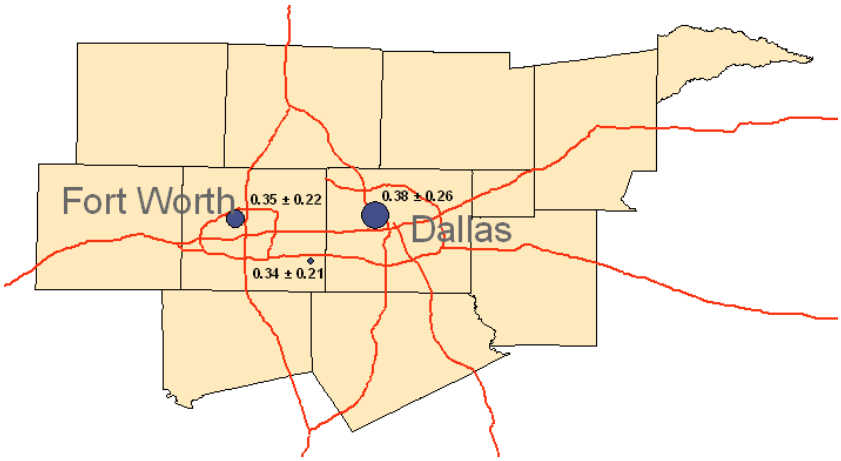


Figure D.3. CO time-series mean (ppm) plotted at monitor locations with graduated symbols for Dallas (a) and Atlanta (b). Labels indicate time-series mean ± time-series standard deviation.

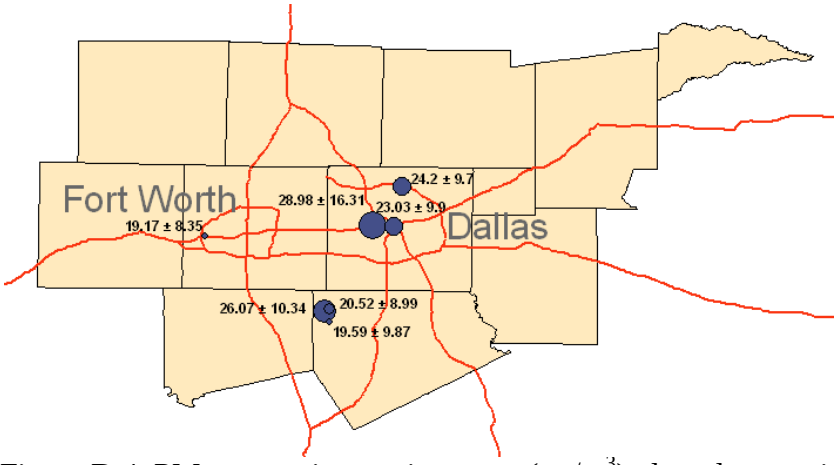


Figure D.4. PM_{10} mass time-series mean ($\mu g/m^3$) plotted at monitor locations with graduated symbols for Dallas (a) and Atlanta (b). Labels indicate time-series mean ± time-series standard deviation.

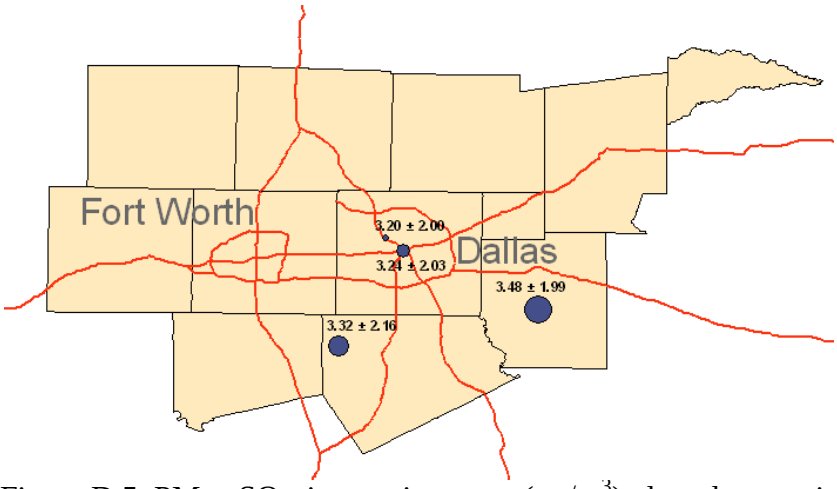


Figure D.5. $PM_{2.5}$-SO_4 time-series mean (μg/m^3) plotted at monitor locations with graduated symbols for Dallas (a) and Atlanta (b). Labels indicate time-series mean ± time-series standard deviation.

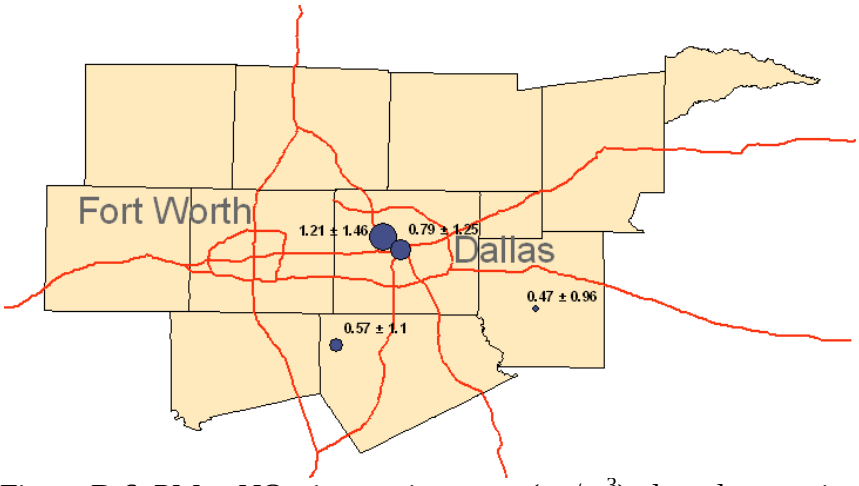


Figure D.6. $PM_{2.5}$-NO_3 time-series mean (μg/m^3) plotted at monitor locations with graduated symbols for Dallas (a) and Atlanta (b). Labels indicate time-series mean ± time-series standard deviation.

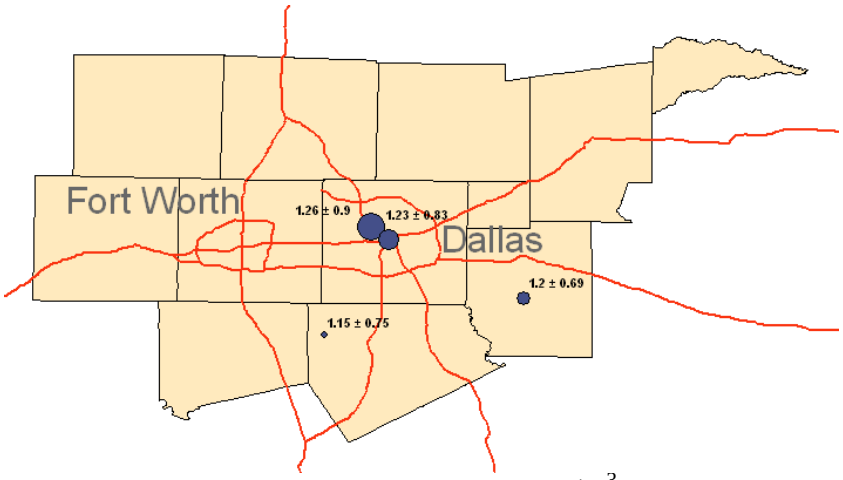


Figure D.7. $PM_{2.5}$-NH_4 time-series mean ($\mu g/m^3$) plotted at monitor locations with graduated symbols for Dallas (a) and Atlanta (b). Labels indicate time-series mean ± time-series standard deviation.

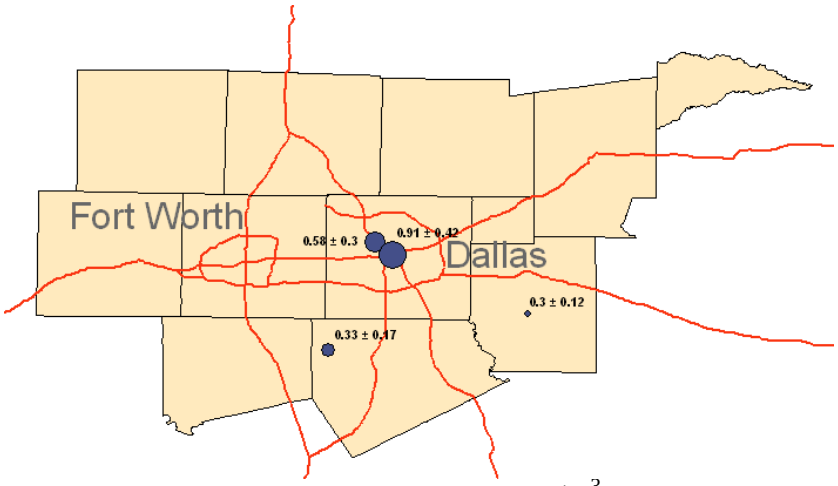


Figure D.8. $PM_{2.5}$-EC time-series mean ($\mu g/m^3$) plotted at monitor locations with graduated symbols for Dallas (a) and Atlanta (b). Labels indicate time-series mean ± time-series standard deviation.

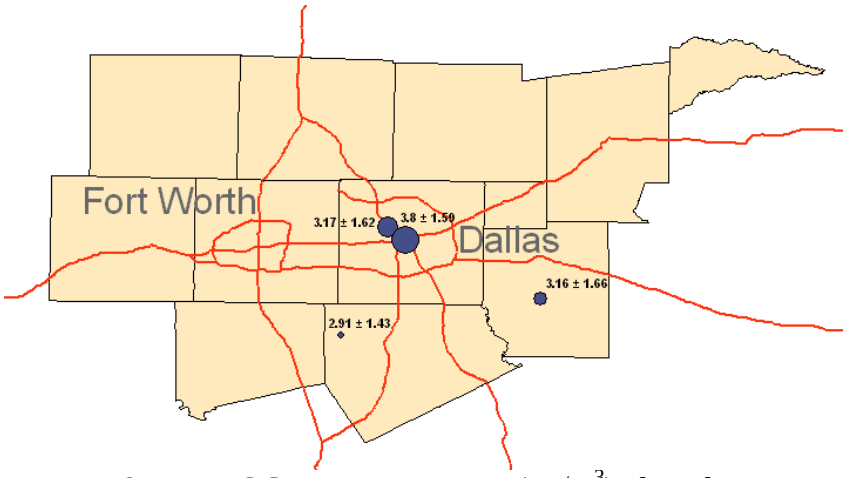


Figure D.9. $PM_{2.5}$-OC time-series mean ($\mu g/m^3$) plotted at monitor locations with graduated symbols for Dallas (a) and Atlanta (b). Labels indicate time-series mean ± time-series standard deviation.

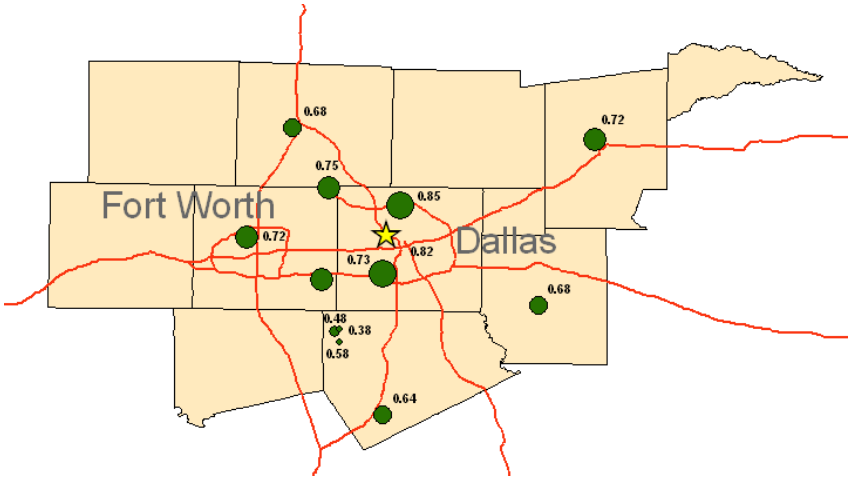


Figure D.10. Correlation of each monitoring site time-series with the central monitor time-series for NO_2 plotted at monitor locations with graduated symbols for Dallas (a) and Atlanta (b). Labels indicate Pearson correlation coefficients and yellow star indicated central monitor location.

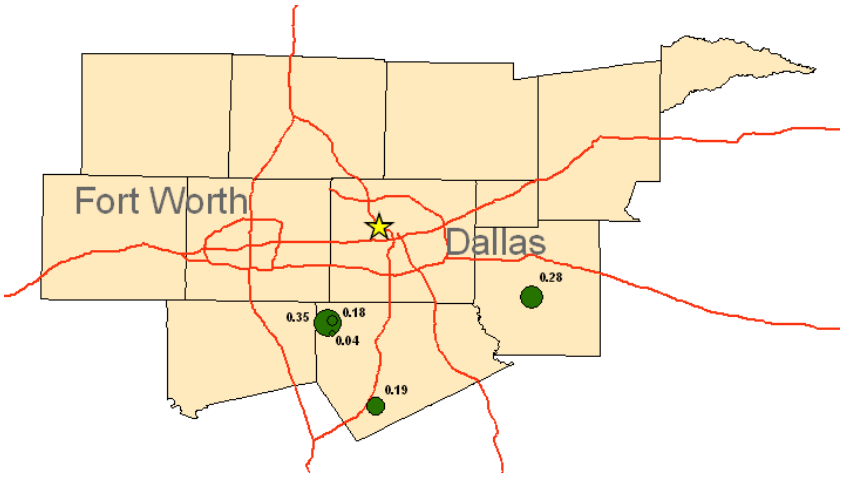


Figure D.11. Correlation of each monitoring site time-series with the central monitor time-series for SO_2 plotted at monitor locations with graduated symbols for Dallas (a) and Atlanta (b). Labels indicate Pearson correlation coefficients and yellow star indicated central monitor location.

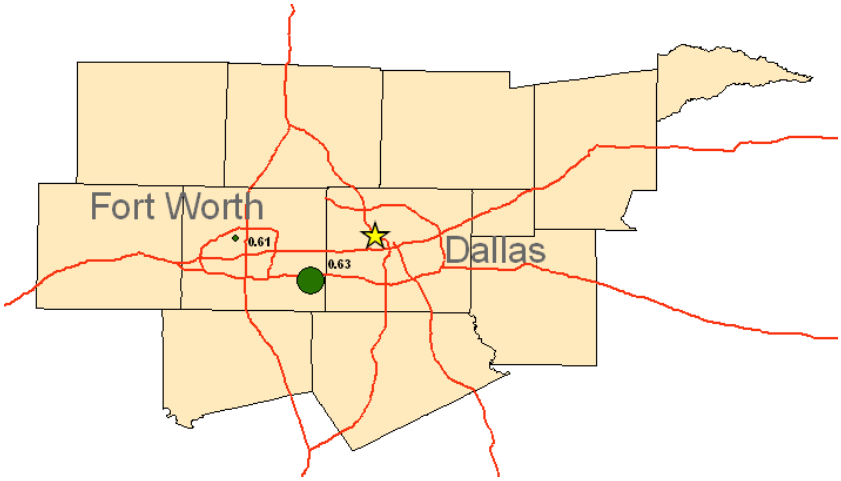


Figure D.12. Correlation of each monitoring site time-series with the central monitor time-series for CO plotted at monitor locations with graduated symbols for Dallas (a) and Atlanta (b). Labels indicate Pearson correlation coefficients and yellow star indicated central monitor location.

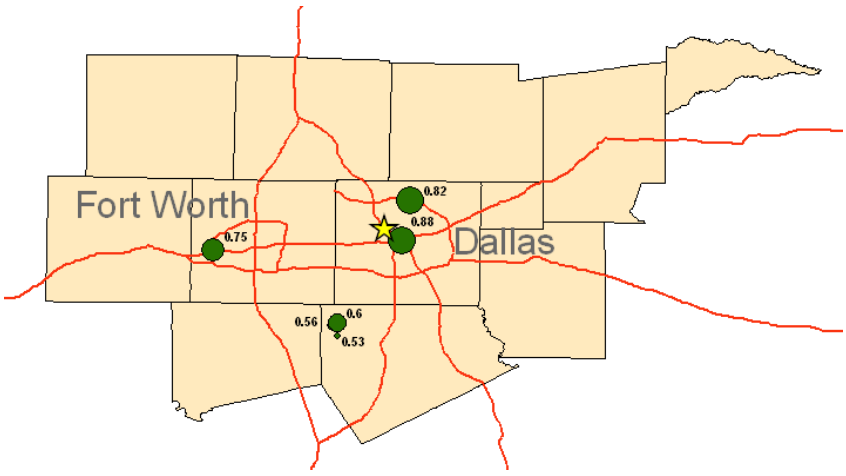


Figure D.13. Correlation of each monitoring site time-series with the South Akard monitoring site time-series for PM_{10} mass (PM_{10} mass measurements were not available at the Hinton Street monitor) plotted at monitor locations with graduated symbols for Dallas (a) and Atlanta (b). Labels indicate Pearson correlation coefficients and yellow star indicated central monitor location.

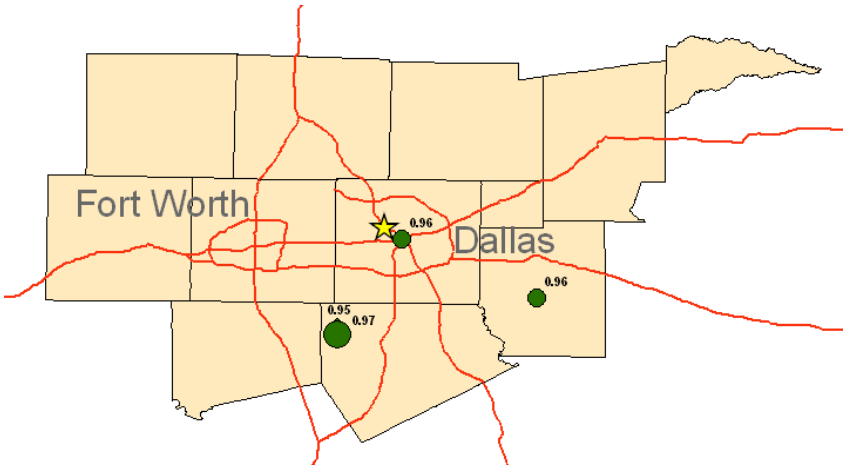


Figure D.14. Correlation of each monitoring site time-series with the central monitor time-series for $PM_{2.5}$-SO_4 plotted at monitor locations with graduated symbols for Dallas (a) and Atlanta (b). Labels indicate Pearson correlation coefficients and yellow star indicated central monitor location.

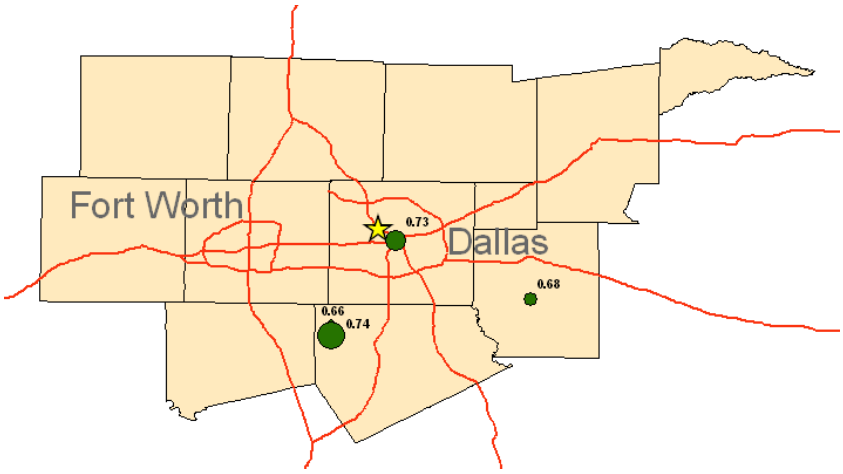


Figure D.15. Correlation of each monitoring site time-series with the central monitor time-series for $PM_{2.5}$-NO_3 plotted at monitor locations with graduated symbols for Dallas (a) and Atlanta (b). Labels indicate Pearson correlation coefficients and yellow star indicated central monitor location.

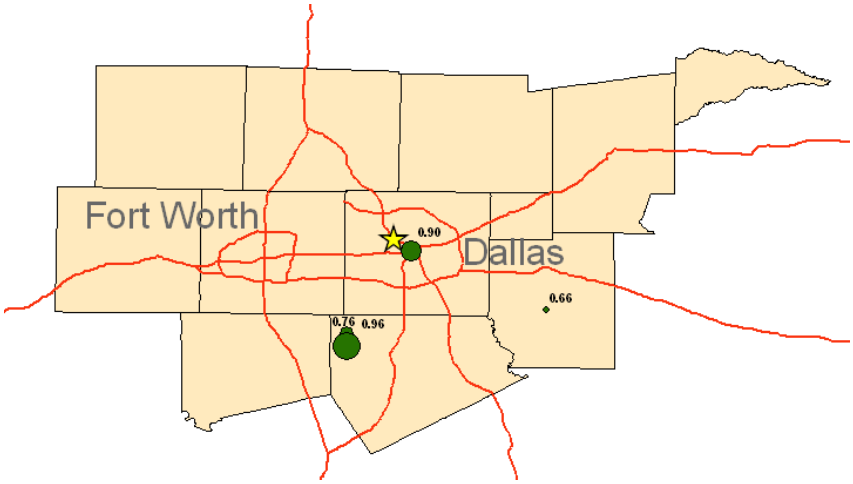


Figure D.16. Correlation of each monitoring site time-series with the central monitor time-series for $PM_{2.5}$-NH_4 plotted at monitor locations with graduated symbols for Dallas (a) and Atlanta (b). Labels indicate Pearson correlation coefficients and yellow star indicated central monitor location.

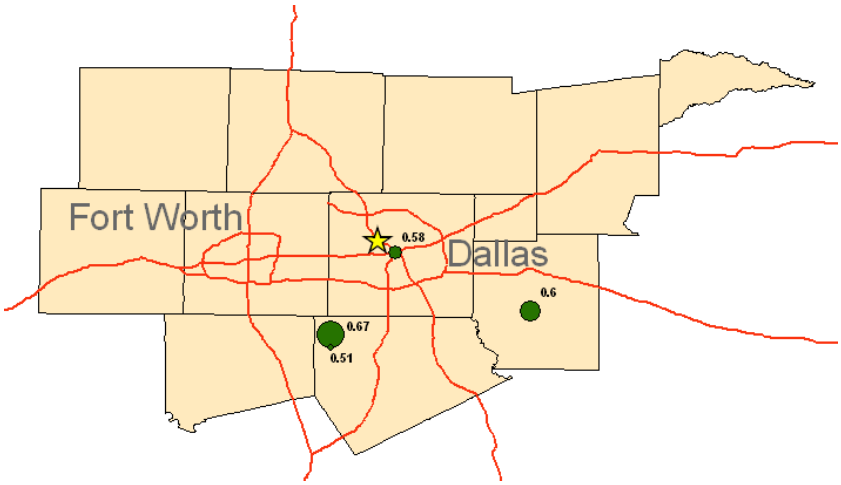


Figure D.17. Correlation of each monitoring site time-series with the central monitor time-series for $PM_{2.5}$-EC plotted at monitor locations with graduated symbols for Dallas (a) and Atlanta (b). Labels indicate Pearson correlation coefficients and yellow star indicated central monitor location.

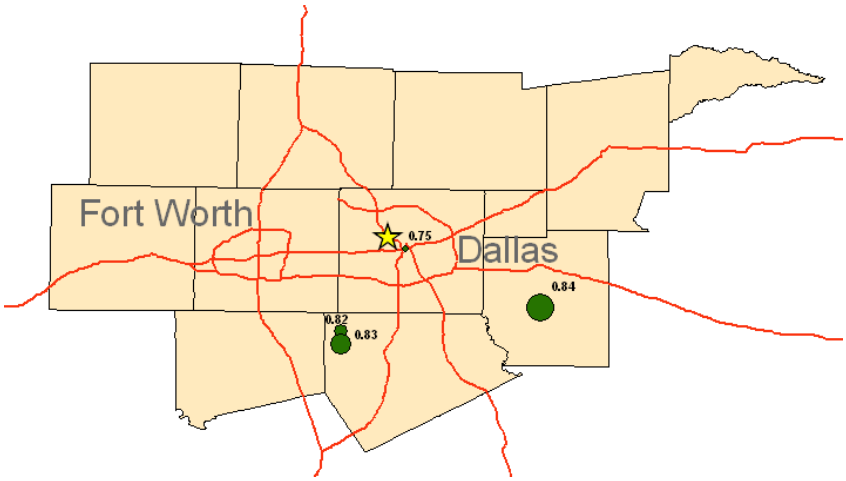


Figure D.18. Correlation of each monitoring site time-series with the central monitor time-series for $PM_{2.5}$-OC plotted at monitor locations with graduated symbols for Dallas (a) and Atlanta (b). Labels indicate Pearson correlation coefficients and yellow star indicated central monitor location.

REFERENCES

Abbey DE, Nishino N, McDonnell WF, Burchette RJ, Knutsen SF, Beeson WL, et al. 1999. Long-term inhalable particles and other air pollutants related to mortality in nonsmokers. American Journal of Respiratory and Critical Care Medicine 159(2): 373-382.

Armstrong BG. 1998. Effect of measurement error on epidemiological studies of environmental and occupational exposures. Occupational and Environmental Medicine 55(10): 651-656.

Brauer M, Hoek G, van Vliet P, Meliefste K, Fischer P, Gehring U, et al. 2003. Estimating long-term average particulate air pollution concentrations: Application of traffic indicators and geographic information systems. Epidemiology 14(2): 228-239.

Brunekreef B, Holgate ST. 2002. Air pollution and health. Lancet 360(9341): 1233-1242.

Butler AJ, Andrew MS, Russell AG. 2003. Daily sampling of PM2.5 in Atlanta: results of the first year of the assessment of spatial aerosol composition in Atlanta study. Journal of Geophysical Research-Atmospheres 108(D1).

Carroll RJ, Ruppert D, Stefanski L. 1995. Measurement Error in Nonlinear Models. London: Chapman & Hall.

Carrothers TJ, Evans JS. 2000. Assessing the impact of differential measurement error on estimates of fine particle mortality. Journal of the Air & Waste Management Association 50(1): 65-74.

Casado LS, Rouhani S, Cardelino CA, Ferrier AJ. 1994. Geostatistical Analysis and Visualization of Hourly Ozone Data. Atmospheric Environment 28(12): 2105-2118.

Chen LP, Mengersen K, Tong SL. 2007. Spatiotemporal relationship between particle air pollution and respiratory emergency hospital admissions in Brisbane, Australia. Science of The Total Environment 373(1): 57-67.

Chow JC, Watson JG, Crow D, Lowenthal DH, Merrifield T. 2001. Comparison of IMPROVE and NIOSH carbon measurements. Aerosol Science and Technology 34(1): 23-34.

Chow JC, Watson JG, Lowenthal DH, Park K, Doraiswamy P, Bowers K, et al. 2008. Continuous and filter-based measurements of PM2.5 nitrate and sulfate at the Fresno Supersite. Environmental Monitoring and Assessment 144(1-3): 179-189.

Cohen AJ, Anderson HR, Ostro B, Pandey KD, Krzyzanowski M, Kunzli N, et al. 2005. The global burden of disease due to outdoor air pollution. Journal of Toxicology and Environmental Health-Part a-Current Issues 68(13-14): 1301-1307.

Darrow L, Klein M, Correa A, Flanders WD, Waller L, Marcus M, et al. 2009. Ambient Air Pollution and Birth Weight in Full-Term Infants in Atlanta, 1994-2004. Epidemiology 20(6): S54-S54.

Dockery DW, Pope CA, Xu XP, Spengler JD, Ware JH, Fay ME, et al. 1993. An Association Between Air Pollution and Mortality in 6 United States Cities. New England Journal of Medicine 329(24): 1753-1759.

Dominici F, Zeger SL, Samet JM. 2000. A measurement error model for time-series studies of air pollution and mortality. Biostat 1(2): 157-175.

Dominici F, Peng RD, Bell ML, Pham L, McDermott A, Zeger SL, et al. 2006. Fine particulate air pollution and hospital admission for cardiovascular and respiratory diseases. Jama-Journal of the American Medical Association 295(10): 1127-1134.

Dutton SJ, Schauer JJ, Vedal S, Hannigan MP. 2009. PM2.5 characterization for time series studies: Pointwise uncertainty estimation and bulk speciation methods applied in Denver. Atmospheric Environment 43(5): 1136-1146.

Edgerton ES, Hartsell BE, Saylor RD, Jansen JJ, Hansen DA, Hidy GM. 2006. The Southeastern Aerosol Research and Characterization Study, part 3: Continuous measurements of fine particulate matter mass and composition. Journal of the Air & Waste Management Association 56(9): 1325-1341.

Fuentes M, Song HR, Ghosh SK, Holland DM, Davis JM. 2006. Spatial association between speciated fine particles and mortality. Biometrics 62(3): 855-863.

Fuller WA. 1987. Measurement Error Models. Chichester: Wiley.

Goldman GT, Mulholland JA, Russell AG, Srivastava A, Strickland MJ, Klein M, et al. 2010. Ambient Air Pollutant Measurement Error: Characterization and Impacts in a Time-Series Epidemiologic Study in Atlanta. Environmental Science & Technology 44(19): 7692-7698.

Goldman GT, Mulholland JA, Russell AG, Strickland MJ, Klein M, Waller LA, et al. 2011. Impact of Exposure Measurement Error in Air Pollution Epidemiology: Effect of Error Type in Time-series Studies. Environmental Health 10: 61.

Gryparis A, Paciorek CJ, Zeka A, Schwartz J, Coull BA. 2009. Measurement error caused by spatial misalignment in environmental epidemiology. Biostatistics 10(2): 258-274.

Hansen DA, Edgerton ES, Hartsell BE, Jansen JJ, Kandasamy N, Hidy GM, et al. 2003. The southeastern aerosol research and characterization study: Part 1-overview. Journal of the Air & Waste Management Association 53(12): 1460-1471.

Hewitt CN. 1991. Spatial Variations in Nitrogen-Dioxide Concentrations in an Urban Area Atmospheric Environment Part B-Urban Atmosphere 25(3): 429-434.

Hinkley D. 1977. On quick choice of power transformation. Applied Statistics 26: 67-69.

Hoek G, Brunekreef B, Fischer P, van Wijnen J. 2001. The association between air pollution and heart failure, arrhythmia, embolism, thrombosis, and other cardiovascular causes of death in a time series study. Epidemiology 12(3): 355-357.

Hogrefe O, Schwab JJ, Drewnick F, Lala GG, Peters S, Demerjian KL, et al. 2004. Semicontinuous PM2.5 sulfate and nitrate measurements at an urban and a rural location in New York: PMTACS-NY summer 2001 and 2002 campaigns. Journal of the Air & Waste Management Association 54(9): 1040-1060.

Hughes JD. 1994. Pan's Travail: Environmental Problems of the Ancient Greeks and Romans. Baltimore: The Johns Hopkins University Press.

Hyslop NP, White WH. 2008. An evaluation of interagency monitoring of protected visual environments (IMPROVE) collocated precision and uncertainty estimates. Atmospheric Environment 42(11): 2691-2705.

International Organization for Standardization. 1995. Guide to the expression of uncertainty in measurement. 1st : Corrected and reprinted, 1995. ed. Genève, Switzerland.

Ivy D, Mulholland JA, Russell AG. 2008. Development of ambient air quality population-weighted metrics for use in time-series health studies. Journal of the Air & Waste Management Association 58(5): 711-720.

Jacobson MZ. 2002. Atmospheric Pollution: History, Science and Regulation. Cambridge: Cambridge University Press.

Jerrett M, Arain A, Kanaroglou P, Beckerman B, Potoglou D, Sahsuvaroglu T, et al. 2005. A review and evaluation of intraurban air pollution exposure models. Journal of Exposure Analysis and Environmental Epidemiology 15(2): 185-204.

Jerrett M, Gale S, Kontgis C. 2010. Spatial Modeling in Environmental and Public Health Research. International Journal of Environmental Research and Public Health 7(4): 1302-1329.

Kaynak B, Hu Y, Martin RV, Sioris CE, Russell AG. 2009. Comparison of weekly cycle of NO2 satellite retrievals and NOx emission inventories for the continental United States. Journal of Geophysical Research-Atmospheres 114.

Krewski D, Rainham D. 2007. Ambient air pollution and population health: Overview. Journal of Toxicology and Environmental Health-Part a-Current Issues 70(3-4): 275-283.

Lee D, Shaddick G. 2010. Spatial Modeling of Air Pollution in Studies of Its Short-Term Health Effects. Biometrics 66(4): 1238-1246.

Lee S, Wang YH, Russell AG. 2010. Assessment of Secondary Organic Carbon in the Southeastern United States: A Review. Journal of the Air & Waste Management Association 60(11): 1282-1292.

Li YH, Guolo A, Hoffman FO, Carroll RJ. 2007. Shared uncertainty in measurement error problems, with application to Nevada test site fallout data. Biometrics 63(4): 1226-1236.

Linn WS, Szlachcic Y, Gong H, Kinney PL, Berhane KT. 2000. Air pollution and daily hospital admissions in metropolitan Los Angeles. Environmental Health Perspectives 108(5): 427-434.

Long RW, McClenny WA. 2006. Laboratory and field evaluation of instrumentation for the semicontinuous determination of particulate nitrate (and other water-soluble particulate components). Journal of the Air & Waste Management Association 56(3): 294-305.

Lowe M. 2007. Sources of Ambient Sufur Dioxide (SO2) in the Metro Atlanta Area. Atlanta: Georgia Institute of Technology.

Mallick B, Hoffman FO, Carroll RJ. 2002. Semiparametric regression modeling with mixtures of Berkson and classical error, with application to fallout from the Nevada test site. Biometrics 58(1): 13-20.

Mar TF, Norris GA, Koenig JQ, Larson TV. 2000. Associations between air pollution and mortality in Phoenix, 1995-1997. Environmental Health Perspectives 108(4): 347-353.

Mendoza-Dominguez A, Russell AG. 2001. Estimation of emission adjustments from the application of four-dimensional data assimilation to photochemical air quality modeling. Atmospheric Environment 35(16): 2879-2894.

Metzger KB, Tolbert PE, Klein M, Peel JL, Flanders WD, Todd K, et al. 2004. Ambient air pollution and cardiovascular emergency department visits. Epidemiology 15(1): 46-56.

Moolgavkar SH, Luebeck EG, Anderson EL. 1997. Air pollution and hospital admissions for respiratory causes in Minneapolis St. Paul and Birmingham. Epidemiology 8(4): 364-370.

Mulholland JA, Butler AJ, Wilkinson JG, Russell AG, Tolbert PE. 1998. Temporal and spatial distributions of ozone in Atlanta: Regulatory and epidemiologic implications. Journal of the Air & Waste Management Association 48(5): 418-426.

National Research Council (U.S.). Committee on Research Priorities for Airborne Particulate Matter. 1998. Research priorities for airborne particulate matter. National Academies Press.

National Research Council (U.S.). Committee on Research Priorities for Airborne Particulate Matter. 2001. Research priorities for airborne particulate matter. III, Early research progress. National Academy Press.

Nunes C, Soares A. 2005. Geostatistical space-time simulation model for air quality prediction. Environmetrics 16(4): 393-404.

Pachon JE, Balachandran S, Hu YT, Weber RJ, Mulholland JA, Russell AG. 2010. Comparison of SOC estimates and uncertainties from aerosol chemical composition and gas phase data in Atlanta. Atmospheric Environment 44(32): 3907-3914.

Peng RD, Bell ML. 2010. Spatial misalignment in time series studies of air pollution and health data. Biostatistics 11(4): 720-740.

Reeves GK, Cox DR, Darby SC, Whitley E. 1998. Some aspects of measurement error in explanatory variables for continuous and binary regression models. Statistics in Medicine 17(19): 2157-2177.

Remy N. S-GEMS: The Stanford geostatistical modeling software: A tool for new algorithms development. In: Proceedings of the Geostatistics Banff 2004: 7th International Geostatistics Conference, Quantitative Geology and Geostatics, 2005. Banff, Vol. 1(Leuangthong O, Deutsch C eds). New York: Springer, 865-871.

Ren C, Tong S. 2008. Health effects of ambient air pollution - recent research development and contemporary methodological challenges. Environmental Health 7.

Sahu S, Mardia K. Recent Trends in Modeling Spatio-Temporal Data. In: Proceedings of the Proceedings of the special meeting on Statistics and Environment organized

by the Societa Italiana di Statistica held in Universita Di Messina, 2005, Vol. 69 83.

Samet JM, Dominici F, Curriero FC, Coursac I, Zeger SL. 2000. Fine particulate air pollution and mortality in 20 US Cities, 1987-1994. New England Journal of Medicine 343(24): 1742-1749.

Sarnat JA, Wilson WE, Strand M, Brook J, Wyzga R, Lumley T. 2007. Panel discussion review: session one - exposure assessment and related errors in air pollution epidemiologic studies. Journal of Exposure Science and Environmental Epidemiology 17: S75-S82.

Sarnat SE, Klein M, Sarnat JA, Flanders WD, Waller LA, Mulholland JA, et al. 2010. An examination of exposure measurement error from air pollutant spatial variability in time-series studies. Journal Of Exposure Science & Environmental Epidemiology 20(2): 135-146.

Sheppard L, Slaughter JC, Schildcrout J, Liu LJS, Lumley T. 2005. Exposure and measurement contributions to estimates of acute air pollution effects. Journal of Exposure Analysis and Environmental Epidemiology 15(4): 366-376.

Soares A. 2001. Direct sequential simulation and cosimulation. Mathematical Geology 33(8): 911-926.

Solomon P, Baumann K, Edgerton E, Tanner R, Eatough D, Modey W, et al. 2003. Comparison of integrated samplers for mass and composition during the 1999 Atlanta Supersites project. Journal of Geophysical Research-Atmospheres 108(D7).

Solomon PA, Chameides W, Weber R, Middlebrook A, Kiang CS, Russell AG, et al. 2003. Overview of the 1999 Atlanta Supersite Project. Journal of Geophysical Research-Atmospheres 108(D7).

Strand M, Vedal S, Rodes C, Dutton SJ, Gelfand EW, Rabinovitch N. 2006. Estimating effects of ambient PM2.5 exposure on health using PM2.5 component measurements and regression calibration. Journal of Exposure Science and Environmental Epidemiology 16(1): 30-38.

US Environmental Protection Agency. 2009. Integrated Science Assessment for Particulate Matter (Final Report). Washington, DC.

Wade KS, Mulholland JA, Marmur A, Russell AG, Hartsell B, Edgerton E, et al. 2006. Effects of instrument precision and spatial variability on the assessment of the temporal variation of ambient air pollution in Atlanta, Georgia. Journal of the Air & Waste Management Association 56(6): 876-888.

White WH, Ashbaugh LL, Hyslop NP, McDade CE. 2005. Estimating measurement uncertainty in an ambient sulfate trend. Atmospheric Environment 39(36): 6857-6867.

Wilson JG, Kingham S, Pearce J, Sturman AP. 2005. A review of intraurban variations in particulate air pollution: Implications for epidemiological research. Atmospheric Environment 39(34): 6444-6462.

Wilson WE, Mar TF, Koenig JQ. 2007. Influence of exposure error and effect modification by socioeconomic status on the association of acute cardiovascular mortality with particulate matter in Phoenix. Journal Of Exposure Science & Environmental Epidemiology 17 Suppl 2: S11-S19.

Zeger SL, Thomas D, Dominici F, Samet JM, Schwartz J, Dockery D, et al. 2000. Exposure measurement error in time-series studies of air pollution: concepts and consequences. Environmental Health Perspectives 108(5): 419-426.

VITA

GRETCHEN TANNER GOLDMAN

Gretchen Tanner Goldman was born and raised in Cheshire, Connecticut, alongside her two younger brothers, Jared and Daniel. Growing up, Gretchen kept busy with academics, gymnastics and varsity track & field. In fall 2002, she enthusiastically entered Cornell University in Ithaca, New York. While at Cornell, Gretchen raced road cycling for the Cornell Cycling Club and was active in the Cornell Chapter of the American Meteorological Society but much of her time was spent doing teambuilding facilitation for Cornell Outdoor Education. This job sparked Gretchen's interest in the natural world and human impacts on it. In particular, she became interested in studying urban air pollution. After received a B.S. in Atmospheric Science from Cornell University in May 2006, she began graduate studies at the Georgia Institute of Technology as a Presidential Fellow. Under the advisement of Drs. James Mulholland and Ted Russell, Gretchen received an M.S. in Environmental Engineering in August 2008 and subsequently pursued a Ph.D. studying air pollution measurement error in time-series epidemiologic studies of acute health outcomes. During her time at Georgia Tech, Gretchen served as President of the Association of Environmental Engineers & Scientists and as a Senator for the Georgia Tech Student Government Association. She was honored as the 2011 Outstanding Ph.D. Candidate for the Environmental Engineering department. Along with her studies, Gretchen is active in environmental advocacy efforts. She served as Chair of the Georgia Tech-wide Bicycle Infrastructure Improvement Committee, which was awarded the 2011 Georgia Tech Environmental Initiative Award and regularly volunteered with the Atlanta Bicycle Coalition and GT Students Organizing for Sustainability's Starter Bikes program, for which she was recognized as Atlanta Bicycle Coalition's 2010 Volunteer of the Year. Gretchen is an avid cyclist and enjoys traveling and spending time outdoors.